

UNIVERSIDAD AUTÓNOMA DE MADRID  
FACULTAD DE CIENCIAS  
DEPARTAMENTO DE BIOLOGÍA MOLECULAR

Diseño y desarrollo funcional de modelos  
vectoriales derivados de lentivirus  
deficientes en integración para la  
modificación génica específica de sitio

TESIS DOCTORAL  
RAÚL TORRES RUIZ  
MADRID, OCTUBRE 2012



La memoria, con título “Diseño y desarrollo funcional de modelos vectoriales derivados de lentivirus deficientes en integración para la modificación génica específica de sitio”, es presentada por D. Raúl Torres Ruiz, licenciado en biología, para optar al grado de doctor en ciencias por la universidad Autónoma de Madrid.

El trabajo experimental presentado en esta memoria ha sido realizado en la unidad técnica de vectores virales, en el centro nacional de investigaciones cardiovasculares (CNIC), bajo la supervisión del Dr. Juan Carlos Ramírez Martínez.

Para que conste a todos los efectos, firma la presente certificación:

VºBº del Director de Tesis

Dr. Juan Carlos Ramírez Martínez





## SUMMARY



Transgenesis procedures in which the position and dosage of the inserted transgene are uncontrolled can lead to unpredictable effects due to disruption of host gene function, compromising data interpretation and placing severe limits on the applicability of this technology. Current methodologies for transgenesis suffer from a strong experimenter effect due to the uncontrolled position and number of inserted transgenes. Therefore, the availability of efficient procedures to generate isogenic cells aimed to obtain comparable results (because of the use of identically modified organisms) could allow to gain more reliable insights into gene function and ultimately, closer to those operating in physiological, pathological, and therapeutic conditions.

We have addressed in this Ph.D. thesis project the development and functional assessment of novel designs aimed at providing solutions to the aforementioned problems. The designs are based on integration-deficient lentivirus (IDLV) and have followed three approaches:

1. Recombinase-mediated cassette exchange (RMCE) platform is based on the integration in the cell genome of an acceptor cassette (genomic landing pad Kas, standing for KO/KI acceptor site) to generate a receptor-modified cell. The second step is the delivery of a DONOR cassette driving the interchange (transfer Trina, standing for *Transfer into acceptor*) delivered by an IDLV. The IDLV encodes in an auto-limited manner the expression of cre recombinase that promotes the RMCE reaction on Kas. A promoter trapping strategy serves to monitor the authentic recombination switching the pattern of expression reporter genes. We have obtained reliable results in both human cell line (HEK293) and primary human cells (BJ fibroblasts) with high efficiency of RMCE (100% of the selected clones). The genomic analysis revealed proper insertion of the donor cassette as well as the absence of integrated copies of the lentivirus, assessed by high sensitivity qPCR. Finally we have demonstrated the versatility and applicability of the platform on episomal plasmids, random integrated copies and, importantly on single copy insertion of Kas in the safe AAVS1 locus.

2.

eGFP-driven fluorescence; excision of floxed genes by incoming cre; and site-specific integration on AAVS1 mediated by rep78.

3. episomes.

We have studied systematically the effect on episomal maintenance of several eukaryotic replication origins (ori) and S/MAR sequences on one-to-one combinations. The total generated bearing those sequences and reporter neoR and eGFP genes have been demonstrated to last the expression of the reporters in the absence of any selection for more than 3 months. The analysis of integration of the constructions by Southern blot on genomic and Hirt's extracts, conventional PCR, and qPCR support the idea of non-integration at the end of the period. Finally, conclusive results were obtained by FISH analysis demonstrating that several combinations allow maintaining the phenotype with no-integrated copies in highly dividing cells (HEK293). Then we have developed novel and efficient episomes supporting partition, i.e. replication in the absence of viral proteins and segregation during karyokinesis.

In summary, we have expanded the current use of vectors exploiting this capabilities to generate a collection of biological tools aimed at directing site specific integration in the genome with minima interference. We can anticipate a wide use in future transgenesis experiments both in research and, hopefully in therapies.



## ÍNDICE



## ABREVIATURAS Y ACRÓNIMOS

1-2

## INTRODUCCIÓN

3-24

1. Modificación génica: la tecnología del DNA recombinante y la biología del tercer milenio
2. Modificación génica: herramientas
3. Modificación génica integrativa
  - 3.1. Integración específica en sitio aceptor natural
    - 3.1.1. Recombinación homologa
    - 3.1.2. Integración natural en sitio específica. Virus AAV
  - 3.2. Integración específica en sitio aceptor artificial.
    - 3.2.1. Nucleasas
    - 3.2.2. Recombinasas. Sistema de recombinación mediado por intercambio de casetes (RMCE)
  - 3.3. Integración al azar en sitio aceptor natural. Lentivirus
    - 3.3.1. Vectores derivados de Retroviridae
    - 3.3.2. Sistemas de generación de partículas lentivirales recombinantes
    - 3.3.3. La integración lentiviral
  - 3.4. Modificación no-integrativa limitada
    - 3.4.1. Vectores lentivirales integrasa deficientes (IDLVs)
  - 3.5. Modificación no-integrativa no-limitada
    - 3.5.1. Episomas replicativos
  - 3.6. ¿Necesitamos siempre transferir genes?. Sistemas de entrega de proteínas
4. Limitaciones y Problemas

## OBJETIVOS

25-26

## RESULTADOS

27-96

MODELO 1. Sistema de intercambio de casete génica mediado por recombinasas (RMCE) y lentivirus deficientes en integración (IDLV)

27-62

1. Objetivo del diseño
2. Elementos del diseño
  - 2.1. RMCE
  - 2.2. Selección de los eventos favorables
  - 2.3. Transferencia lentiviral
  - 2.4. Expresión limitada
  - 2.5. Descripción funcional del diseño
    - 2.5.1. Diseño del casete aceptor: KAS1.0

- 2.5.2. Diseño del casete donador: TRINA1.0
- 2.5.3. Generación de los elementos génicos y biológicos
- 3. Prueba de función 1: RMCE en copias episomales
  - 3.1. Análisis de eventos positivos de recombinación por resistencia
  - 3.2. Análisis de eventos positivos por fluorescencia
- 4. Prueba de función 2.: Sistema de copias integradas
  - 4.1. Sistema de entrega lentiviral. Generación y características
    - 4.1.1. Caracterización del vector lentiviral integrasa deficiente (IDLV)
  - 4.2. Generación de la línea estable HEK293 KAS
    - 4.2.1. GLP (genetic landing pad): Sustrato cromosómico
  - 4.3. RMCE mediado por IDLV
    - 4.3.1. Caracterización genotípica de células positivas tras la reacción de RMCE
    - 4.3.2. Optimización del proceso de RMCE mediado por IDLV
    - 4.3.3. Verificación de la expresión limitada de cre
- 5. Prueba de funcionalidad 3: RMCE sobre copia única integrada en sitio específica en HEK293A
  - 5.1. Generación del GLP en un locus determinado
    - 5.1.1. Transferencia por recombinación homóloga. Locus humano ROSA26
    - 5.1.2. Transferencia por recombinación homóloga con ZFN. Locus AAVS1
      - 5.1.2.1. Caracterización genotípica de clones 293A KAS2.0
        - a) Análisis mediante PCR
        - b) Análisis mediante FISH
      - 5.1.2.2. Caracterización fenotípica de clones 293A KAS2.0
        - a) Ensayo de clonogenicidad
        - b) Ensayo de estabilidad
  - 5.2. Generación de clones por RMCE
    - 5.2.1 Caracterización fenotípica
    - 5.2.2 Caracterización genotípica
      - a) Análisis por PCR
      - b) Análisis por secuenciación
      - c) Análisis mediante qPCR
  - 5.3. Optimización del sistema mediante el uso de dos IDLVs
- 6. Prueba de función 4. Versatilidad del sistema
  - 6.1. Recombinación homóloga en línea celular primaria (Fibroblastos BJ)
    - 6.1.1 Caracterización genotípica de clones BJ KAS2.0
      - a) Análisis mediante PCR
    - 6.1.2 Caracterización fenotípica de clones BJ KAS2.0
      - a) Ensayo de estabilidad
  - 6.2. Generación de clones BJ por RMCE
    - 6.2.1 Caracterización fenotípica
    - 6.2.2 Caracterización genotípica



## MODELO 2. Sistema de entrega de enzimas de recombinación

. Aplicación a la integración sitio específica mediada por rep78 de AAV para inserción en el locus humano AAVS1. **63-78**

1. Objetivo del diseño
2. Elementos del diseño
  - 2.1. Fusiones proteicas
  - 2.2. Proteínas testadas
3. Estabilidad y producción de la partículas
  - 3.1. Producción de particulares de entrega de eGFP
  - 3.2. Transducción con partículas eGFP
4. Funcionalidad de la recombinasa cre
  - 4.1. Generación y verificación de un sistema reportero
  - 4.2. Estudio funcional de la proteína cre transducida
5. Aplicación a la integración dirigida en AAVS1
  - 5.1. Generación de un sistema reportero
  - 5.2. Prueba de función (I). Transfección transitoria con plásmidos
  - 5.3. Prueba de función (II): Transducción con partículas
  - 5.4. Prueba de función (y III): Transducción con partículas
  - 5.5. Caracterización genotípica. Análisis mediante Southern blot

## MODELO 3.

mantenimiento y partición.

entrega de formas extracromosómicas con

**79-96**

1. Objetivo del diseño
2. Elementos del diseño.
  - 2.1.
  - 2.2. Orígenes de replicación
  - 2.3. S/MAR
  - 2.4. Descripción del diseño
3. Estabilidad y viabilidad episomal
  - 3.1. Análisis estructural
  - 3.2. Producción, transducción y cinética de la estabilidad
4. Caracterización genotípica
  - 4.1. Caracterización genotípica por qPCR
  - 4.2. Análisis mediante PCR convencional
  - 4.3. Detección episomal mediante FISH

## **DISCUSIÓN**

97-112

1. RMCE
  - 1.1. Funcionalidad del sistema. Copias episomales
  - 1.2. Copias integradas al azar
  - 1.3. Copia única integrada en sitio específico
2. Sistema de entrega de proteínas con interés biológico
  - 2.1. Aplicación al desarrollo de híbridos
3. Episomas replicativos
4. Futuras aproximaciones

## **CONCLUSIONES**

113-114

## **MATERIAL Y MÉTODOS**

115-126

1. Cultivos celulares
2. Generación de plásmidos
3. Transfecciones celulares
4. Lentivirus recombinantes
5. Selección de clones celulares
6. Caracterización genética de clones celulares
7. Caracterización fenotípica

## **BIBLIOGRAFÍA**

127-140

## **ANEXO I. Tabla de primers o cebadores**

## **ABREVIATURAS Y ACRÓNIMOS**



|                 |   |
|-----------------|---|
| AAV             | Virus Adeno-asociado  |
| Ab              | Anticuerpo  |
| Ad              | Adenovirus  |
| BAC             | <i>Bacterial artificial chromosome</i> Cromosoma artificial de bacterias              |
| CA              | proteína de la cápsida lentiviral   |
| CAG             | Promotor híbrido de CMV y beta actina de pollo  |
| cDNA            | <i>complementary deoxiribonucleic acid</i> Acido desoxirribonucleico complementario   |
| ChFP            | <i>Cherry Fluorescent protein</i> , Proteína Roja Fluorescente                        |
| cm <sup>2</sup> | centímetro cuadrado   |
| CMV             | Citomegalovirus, hace referencia al promotor  |
| DAPI            | 4',6-diamidino-2-fenilindol   |
| DNA             | <i>deoxiribonucleic acid</i> Acido desoxirribonucleico                                |
| dNTPs           | Desoxinucleótidostrifosfato   |
| DSB             | <i>Double strand break</i> , rotura de doble hebra del DNA                            |
| EDTA            | Ethylen-Diamin Tetra-Acetao   |
| EE              | Estabilización episomal   |
| EF1a            | Factor de elongación 1, hace referencia al promotor                                   |
| eGFP            | <i>Enhanced Green Fluorescent Protein</i> Proteína Verde Fluorescente mejorada        |
| ETF             | Effective targeting frequency   |
| FBS             | Suero bovino fetal  |
| FISH            | <i>Fluorescence in situ hybridization</i> Hibridación de fluorescencia <i>in situ</i> |
| GLP             | <i>Genetic landing pad</i> , Lugar de emplazamiento genómico                          |
| GOI             | <i>Gene of interest</i> , gen de interés  |
| HAC             | <i>Human artificial chromosomes</i> , cromosomas artificiales humanos                 |
| HBSS2x          | <i>Hanks Balance solution 2x</i>  |
| HEK             | <i>Human embryonic kidney</i> , Riñón Embrionario Humano                              |
| hESC            | <i>human Embryonic Stem Cell</i> , células troncales humanas de origen embrionario    |
| HIV             | <i>Human immunodeficiency virus</i> , Virus de la Inmunodeficiencia Humana            |
| HR              | <i>Homologous recombination</i> , recombinación homóloga                              |
| IDLV            | <i>Integrase Deficient Lentivirus</i> , Lentivirus defectivos en integración          |
| iPSC            | <i>Induced pluripotent stem cells</i> , células pluripotentes inducidas               |
| IRES            | <i>Internal ribosome entry site</i> , sitio interno de entrada del ribosoma           |
| ITR             | <i>Inverted terminal repeats</i> , Repeticiones Terminales Invertidas                 |
| KAS             | <i>Knock-in/out acceptor site</i> , emplazamiento genómico acceptor                   |
| Kb              | Kilobases   |
| KDa             | Kilodalton  |
| KO              | <i>Knock-Out</i>  |
| LS              | Lentivirus episomales replicativos o lentisomas                                       |
| LTR             | <i>Long terminal repeats</i> , terminaciones largas repetidas                         |
| LV              | Lentivirus  |
| MA              | proteína de la matriz lentiviral  |
| MCS             | <i>Multicloning site</i> , Sitio de Clonación Múltiple                                |
| MFI             | <i>Mean fluorescence intensity</i> , intensidad media de fluorescencia                |
| MLV             | Moloney Leukemia Virus  |

|                 |  |
|-----------------|--|
| MOI             | <i>Multiplicity of infection</i> , Multiplicidad de Infección                                    |
| mRNA            | <i>Messenger RNA</i> , ácido desoxiribonucleico mensajero  |
| NC              | proteína de la nucleocápsida lentiviral  |
| NHEJ            | <i>Non homologous end joining</i> , unión no homóloga de extremos                                |
| OpCC            | <i>Once per cell cycle</i> , una vez por ciclo celular   |
| ORF             | <i>OpenReading Frame</i> , Marco abierto de lectura  |
| Oris            | orígenes de replicación  |
|                 |  |
| pb              | Pares de base  |
| PBS             | <i>Phosphate saline buffer</i> , solución tamponadora de fosfato salino                          |
| PCR             | <i>Polymerase chain reaction</i> reacción en cadena de la polimerasa                             |
| PFA             | Paraformaldehído   |
| PGK             | <i>Phosphoglycerate kinase</i> , fosfoglicerato quinasa, hace referencia a un promotor           |
| PIC             | <i>Pre Integration complex</i> , complejo de pre-integración                                     |
| polyA           | Secuencia de Poli-Adenilación  |
| qPCR            | <i>quantitative PCR</i> PCR cuantitativa   |
| RBS             | <i>Rep binding site</i> , Sitio de unión de Rep  |
| RMCE            | <i>Recombinase mediated cassette exchange</i> , recombinación mediada por intercambio de casetes |
| RNA             | <i>ribonucleic acid</i> ácido ribonucleico   |
| rpm             | Revoluciones por minuto  |
| RT              | Retrotranscriptasa   |
| Sb              | <i>Southern blot</i>   |
| SDS             | Duodecil sulfato de sodio  |
| SIDA            | Síndrome de inmunodeficiencia adquirida  |
| SIDD            | <i>Stress induced dúplex destabilized</i> , inducción de la desestabilización del DNA            |
|                 |  |
| SV40            | <i>Simian Vacuolating virus 40</i> , Virus de Simio 40   |
| S/MAR           | <i>Scaffold/Matrix attachment regions</i> , regiones de unión a la matriz                        |
| TRINA           | <i>Transfer into acceptor site</i> , vector de transferencia en el sitio aceptor                 |
| trs             | <i>Terminal resolution site</i> , sitio de resolución terminal                                   |
| TU              | <i>Transduction units</i> , unidades de transducción   |
| UV              | Ultravioleta   |
| VLP             | <i>Virus like particles</i> , partículas similares a virus                                       |
| VSV-G           | <i>G-protein Vesicular stomatitis virus</i>  |
| ZFN             | <i>ZincFinger nucleases</i> , nucleasas de dedos de Zinc   |
|                 |  |
| <sup>32</sup> P | Radio isótopo de fósforo 32  |
| °C              | Grados Celsius   |
| ΔG°             | Variación en la energía libre  |
| μg              | Microgramo   |
| μl              | Microlitro   |
| μM              | Micromolar   |
| Ψ               | Señal de empaquetamiento de virus  |

## INTRODUCCIÓN





En esencia, los procesos reproductivos no partenogenéticos, son nada más (y nada menos) que la combinación de genomas de distintos individuos. La observación de que en el intercambio de genes se perpetúan los caracteres, forma parte de la cultura ancestral del hombre. Aún así, su comprensión tuvo que esperar hasta el siglo XIX. La selección de organismos vivos por la posesión de unas características concretas no es un hecho reciente, sino que se lleva practicando siglos en la transformación de plantas y animales, con el fin de obtener mejores cosechas, o para domesticar animales salvajes. La conquista de la base molecular de los genes permitió avanzar en la idea de que el material génico pudiera ser manipulado y modificado, para así ser transferido entre células y organismos, con el fin de obtener caracteres fenotípicos concretos. El desarrollo de dicha tecnología ha permitido el nacimiento de disciplinas como la biología molecular y celular (tal como las entendemos hoy día), y la medicina genómica diagnóstica y terapéutica. A pesar de ello quedan muchas preguntas por contestar y suficientes avances básicos que desarrollar, para llevar la terapia con genes al escenario al que pertenece la ingeniería, el de la aplicación<sup>1</sup>.

## **1. Modificación génica: la tecnología del DNA recombinante y la biología del tercer milenio**

Una definición del término ingeniería genética es “la modificación directa y deliberada de las características de un organismo mediante la manipulación de su material genético” [1]. En base a ello, se podría datar la aparición de la misma en la década de los 70 del siglo pasado, aunque el término fue acuñado por vez primera en el año 1932, en el sexto congreso internacional de genética [1], antes incluso de los experimentos de Avery, MacLeod y McCarthy [2] e incluso con más de dos décadas de antelación al descubrimiento de la doble hélice de DNA [3].

Se trata de un término que ha dado lugar a toda una disciplina de la biología, gracias a la cual se ha posibilitado en apenas 40 años la generación de imponderables avances en la comprensión de los sistemas biológicos, constituyendo un paradigma denominado en ocasiones la revolución biológica. Se ha elaborado una tabla resumen (tabla 1) (sabiendo que es incompleta) con aquellos hitos que personal y subjetivamente consideramos más relevantes, para facilitar la visualización de la celeridad con que se ha producido el avance en el desarrollo de las herramientas; que en conjunto, dieron lugar al nacimiento (científico) de la ingeniería genética. Dicha revolución surge de la aplicación de los avances genéticos y moleculares al ámbito biotecnológico (farmacéutico, industrial) [4-6] y a la investigación y medicina [7-10] o incluso al servicio más lúdico y frívolo [11].

---

<sup>1</sup> Dada la necesaria limitación de espacio y de concreción expositiva, en la introducción de esta memoria se describen algunos de los métodos genéticos que se han aplicado con éxito a células humanas y no humanas para su modificación genética. Se ha evitado explicar/comparar/discutir modelos, sistemas y metodologías que no se han empleado a lo largo del trabajo experimental, si bien y aún a riesgo de omisiones imperdonables, aquellas alternativas se mencionan y referencian para permitir al interesado profundizar en dichos aspectos.

| Hito   | Año  | Autores                               |
|--|------|---------------------------------------|
| Descubrimiento de la doble hélice de DNA   | 1953 | J Watson, F Crick & R Franklin        |
| Desarrollo de la primera línea celular humana. HeLa  | 1953 | GO Gey & H Lacks                      |
| Aislamiento de la DNA polimerasa   | 1958 | Coenberg                              |
| Descubrimiento de elementos reguladores en el genoma. Operón   | 1959 | F Jacob & J Monod                     |
| Descifrando el código genético   | 1966 | M Nirenberg, H Mathaei & Severo Ochoa |
| Descubrimiento de la DNA Ligasa  | 1967 | M Gellert                             |
| Descubrimiento de las enzimas de restricción   | 1968 | S Linn & W Arber                      |
| Transferencia génica in vitro mediante SV40  | 1968 | R Dulbecco                            |
| Aislamiento y descripción de la transcriptasa reversa. Cambio del Dogma Central de la Biología Molecular | 1970 | D Baltimore, H Temin & S Mizutani     |
| Transformación bacteriana mediante plásmidos   | 1972 | S Cohen                               |
| Generación in vitro de la primera molécula de DNA recombinante   | 1972 | P Berg                                |
| Introducción de DNA en células eucariotas mediante transfección transitoria                              | 1973 | Graham & van der Eb                   |
| Construcción del primer plásmido recombinante y su introducción en un organismo                          | 1973 | S Cohen & P Boyer                     |
| Primer intento, in vitro, de corregir una enfermedad genética mediante el uso de virus                   | 1973 | S Rogers                              |
| Primer animal modificado genéticamente   | 1974 | R Jaenisch & B Mintz                  |
| Secuenciación del DNA  | 1977 | W Gilbert, A Maxam & F Sanger         |
| Construcción de los primeros vectores retrovirales   | 1981 | Shimotohno, Temin, Wei & Tabin        |
| VIH como causante del SIDA   | 1983 | L Montagnier & Gallo                  |
| Amplificación en cadena mediada por polimerasa (PCR)   | 1985 | K Mullis                              |
| Primer ensayo clínico de terapia génica  | 1990 | French Anderson                       |

**Tabla 1. Descubrimientos relevantes en el ámbito de la modificación génica.** Principales hitos, elegidos subjetivamente, que han conducido al desarrollo y establecimiento de la ingeniería genética. Se ha representado tanto el descubrimiento, como el año en que se llevó a cabo y el autor o autores que lo realizaron.

El desarrollo pionero de la genética como el de los trabajos con *Drosophila sp.* de Thomas H. Morgan o con *E. coli* de M. Delbruck se realizaron mediante aproximaciones de genética directa (del fenotipo al gen). La genética reversa (del gen a la función) aunque están en la base de los experimentos de S. Benzer con fagos y hongos, no pudieron desarrollarse hasta el surgimiento de la tecnología del DNA recombinante [12]. De su desarrollo y del de las técnicas de transferencia de genes (ver más abajo) surge la manera actual de indagar en la función génica: la ganancia y pérdida de función (GoF y LoF del inglés) mediada por la transferencia manipulada o no de elementos génicos, en células y organismos. La extensión de dichos procedimientos con fines terapéuticos es la terapia génica: el tratamiento de patologías humanas, mediante la entrega de material génico, con el fin de poder reemplazar o corregir copias erróneas de los genes causantes [13].

La primera vez que se propuso tal aproximación (hipótesis entonces) fue en 1972, a lo que se llegó por la combinación en el conocimiento de la base molecular de determinadas patologías y el desarrollo realizado en vectores de entrega de DNA [14]. Desde entonces el desarrollo de herramientas ha permitido el avance, no sin antes solventar numerosos inconvenientes a lo largo de todos estos años, en gran número de ensayos clínicos que tienen como base la transferencia génica (1843 ensayos clínicos, <http://www.abedia.com/wiley/phases.php>).<sup>2</sup>

Una frontera científica en biología médica se ha abierto con la disponibilidad de células madre embrionarias humanas [21] y el desarrollo de la tecnología que permite generar células pluripotentes (iPSC) [22]. La eficiencia en el desarrollo de las aplicaciones de transferencia de genes en estudios de diferenciación y pluripotencia están en la base de dichas tecnologías; ofreciendo la oportunidad de estudiar los mecanismos moleculares y celulares de enfermedades tanto monogénicas, como multifactoriales, en tipos celulares relevantes, para poder identificar tratamientos alternativos [23].

Para que este concepto, denominado “enfermedad en una placa de cultivo” sea exitoso es esencial establecer sistemas experimentales en los que la única variable modificada, sea la lesión genética causante de la enfermedad [23]. La óptima diferenciación de células hESCs o iPSCs puede generar modelos de organogénesis, producir células y tejidos para terapias o identificar componentes relevantes para el tratamiento de enfermedades. El marcado genético permitirá el análisis de alto rendimiento (*high-throughput screening*) para identificar moléculas implicadas en especificación de linajes; y la expresión ectópica o el silenciamiento de genes, escrudiñar genes clave implicados en la diferenciación hacia linajes específicos o el trazado celular, por citar algunos ejemplos.

---

<sup>2</sup> La falta de conocimiento condujo a errores en los ensayos iniciales; en unos casos debido al agente terapéutico utilizado [15-17], y en otros por la manera en que se llevó a cabo el ensayo [18]. E incluso más recientemente, debido a los efectos a largo plazo que pueden implicar los vectores en el genoma u organismo diana [19] y su interacción con el sistema inmune [20].

## 2. Modificación génica: herramientas

El desarrollo de vectores de transferencia génica eficaces es uno de los factores clave que determinan la aplicabilidad de los mismos en procesos de transgénesis. La elección del vector a usar en los protocolos de transferencia génica es de vital importancia, tanto en lo relativo a los niveles de bioseguridad requeridos, como en lo que concierne a su entrega o vía de administración; además se ha de considerar su nivel de expresión, la estabilidad que confiere, la retención a lo largo del tiempo y sus posibles efectos adversos (o preferiblemente, la ausencia de ellos) [24].

Conceptualmente se han desarrollado dos tipos de vectores, o dos grandes grupos con un origen diferenciado, los vectores virales y los no virales. Dentro de los no virales, los transgenes pueden introducirse mediante procesos químicos, como lípidos catiónicos o polímeros, o por métodos físicos, como bombardeo de DNA, microinyección, electroporación, magnetofección o ultrasonidos [25].

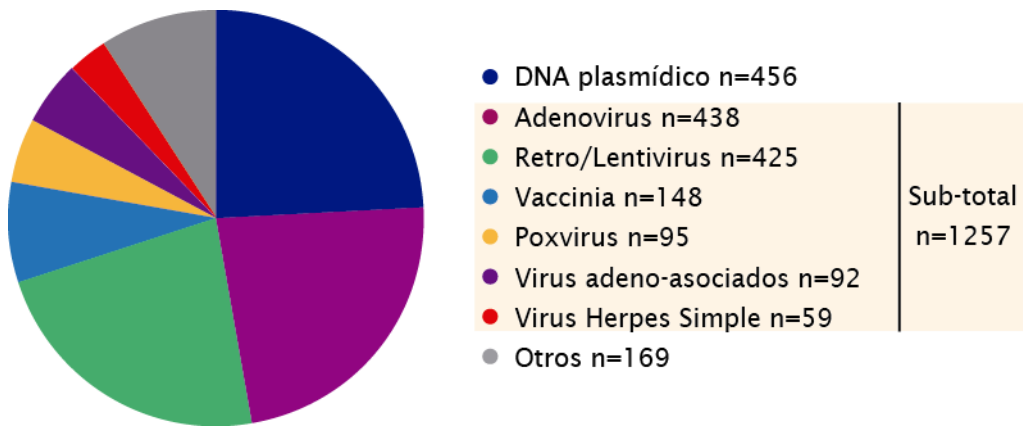
Hasta la fecha pocos estudios se han presentado que hagan comparaciones directas de todos éstos métodos. En realidad se recomienda que se comparen en cada laboratorio múltiples reactivos para encontrar el más efectivo conforme a los propósitos particulares (tipo celular, sensibilidad celular, *in vivo/in vitro*, etc). Se han comparado nucleofección, electroporación y lipofección [26], [27]; o componentes químicos como FuGENE 6, Lipofectamine 2000 (basados en lípidos) y ExGen 500 (basado en reactivos poliméricos-catiónicos) o entre ExGen 500 y GeneJammer [27], [28]. De entre todos ellos el mejor posicionado, realizando una comparativa global de todos los artículos referenciados es el que emplea GeneJammer. No obstante las diferencias en tipo celular y condiciones de cultivo han de considerarse, para cada caso concreto [29]; además, el problema general asociado a este tipo de metodología reside principalmente en la baja eficiencia inherente cuando dichos vectores son extrapolados desde sistemas *in vitro* a organismos completos [30].

Dado que el trabajo experimental de esta memoria se ha realizado con vectores recombinantes derivados de virus, centraremos la atención en las partículas virales recombinantes, que consiguen solventar algunos problemas con una eficiencia relativamente alta y por tanto con una mayor probabilidad de éxito. Ello se debe al hecho de que durante el proceso de evolución los virus han desarrollado interacciones específicas con los receptores presentes en la membrana celular [31], para de esta manera poder internalizarse eficazmente; incluso evadiendo los mecanismos celulares de degradación [32]. Esto los convierte en herramientas moleculares idóneas para ser aplicadas en procedimientos de transferencia génica [33].

Los vectores recombinantes derivados de virus se han desarrollado a partir del concepto, experimentalmente comprobado, de que la transfección de genomas en sistemas celulares permisivos (células productoras) genera viriones [34]. La segregación de los componentes génicos virales en tres grupos: (i) secuencias de reconocimiento de genomas o de

empaquetamiento, generalmente flanqueantes y compuestas de secuencias terminales repetidas [35] , (ii) ORFs no estructurales y (iii) ORFs estructurales, y la constatación de que las primeras funcionan en genomas heterólogos y de que las dos últimas son funcionales en *trans*, permitieron el diseño de estrategias de generación de vectores virales recombinantes. Éstos de basan en tecnología de DNA recombinante, mediante su aporte a través de un sistema de plásmidos, su incorporación en células productoras y/o empaquetadoras [33], junto con plásmidos susceptibles de ser empaquetados. Veremos en mas detalle este punto a continuación para el sistema derivado de lentivirus (LV). Otra de las características más comunes para los sistemas virales de transferencia génica es que son no-replicativos por lo que los componentes virales aportados en *trans* están en construcciones sin secuencias de empaquetamiento, y recíprocamente, las secuencias empaquetadas aportadas en el plásmido lanzadera, (*shuttle o transfer*, según autores) contienen el menor número de secuencias virales posibles [33].

Hay un gran número de familias virales de las que se han derivado dichos vectores si bien en ensayos clínicos de terapia génica los más utilizados son sistemas basados en (1) virus de genoma de DNA: *Adenoviridae* (Mastadenovirus, adenovirus, Ad), *Poxviridae* (Orthopoxvirus, vaccinia virus, VACV; Avipoxvirus, canarypox, CNPV), *Parvoviridae* (Dependovirus, virus adeno-asociado, AAV) y *Herpesviridae* (Alphaherpesvinae, virus del herpes simple, HHV); o (2) derivados de virus con genoma de RNA: *Retroviridae* (Gammaretrovirus, MLV; Lentivirus, LV). En total suponen entorno al 68% de los vectores empleados en ensayos clínicos (figura 1). Estos sistemas virales poseen una gran variedad de ventajas que les hace idóneos para su empleo en ensayos de terapia génica [36].



**Figura 1. Ensayos clínicos de terapia génica.** Número de ensayos clínicos de terapia génica organizados según tipo de vector utilizado. Está representado tanto el porcentaje, como el número de ellos. Se han enmarcado los correspondientes a los vectores virales recombinantes.

Queda fuera del objetivo de esta tesis el referir tanto las características, como las ventajas e inconvenientes de empleo de cada uno de ellos por lo que se referencian alguna excelentes revisiones [37-40]. Dichas propiedades específicas harán que el vector idóneo varíe según los requerimientos que exija su aplicación. El trabajo experimental de esta tesis se ha realizado con vectores derivados de lentivirus, en concreto, del virus de la

inmunodeficiencia humana tipo 1 (HIV-1) sobre los que mencionaremos algunas propiedades que consideramos han servido para el diseño de los objetivos de la misma.

### 3. Modificación génica integrativa

El destino intracelular de las construcciones génicas transferidas puede ser la **integración** en el genoma residente de la célula o la permanencia como forma extra-cromosómica denominada **episomal o no-integrada**.

Dentro de otras consideraciones se puede contemplar, la integración:

- A. en sitio aceptor natural, es decir, existente en la célula parental, sin modificaciones previas, o
- B. sobre sitios aceptores artificiales, por cuanto la célula diana ha de generarse previamente para contener un sitio, secuencia o locus sobre el que se dirija la integración (GLP, del inglés *genetic landing pad*)

Por último, la integración puede producirse sobre un locus o secuencia específica, que se denomina **integración en sitio-específico** (SSI, del inglés) o integrarse sin preferencia de secuencia, **integración al azar** aunque ello no signifique equivalencia a lo largo del genoma aceptor, puesto que se han descrito integraciones preferenciales en regiones determinadas (*hot spots*) con diversos vectores virales[41], [42].

#### 3.1. Integración específica en sitio aceptor natural

##### 3.1.1 Recombinación homologa

Basado en procesos celulares de reparación del DNA, y mediante el empleo de secuencias homólogas es posible llevar a cabo la integración sitio específica en un locus preciso del genoma[9]. El sistema empleado exhaustivamente en sistemas múridos ha permitido la eliminación de genes utilizando ratones *knock-out* (KO) o la introducción mediante el empleo de *loci* específicos (ROSA26, HPRT, ENY, CCR5). Esta aproximación requiere de una construcción con brazos de homología de longitud variable, entre 5-12 kb flanqueando la casete de interés[43]. La frecuencia es baja, entre  $10^{-5}$  y  $10^{-8}$  [44], excepto para *loci* específicos como ROSA26 y HPRT, lo que obliga al análisis de un número elevado de clones, tanto por PCR como mediante *Southern blot*; además de la utilización de sistemas de selección positiva y negativa. Todo ello con el fin de poder caracterizar correctamente los clones celulares y minimizar la obtención de falsos positivos [45], [46].

El problema que conlleva este método es la falta de versatilidad y eficiencia y que sus tasas de especificidad son demasiado bajas [47]; además, se necesitan fragmentos de homología bastante amplios, para poder llevar a cabo el proceso, lo que limita su utilización con

sistemas de entrega virales [48], [49]. Por tanto se hace necesario el empleo y desarrollo de nuevas herramientas que puedan solventar estos problemas.

### 3.1.2 Integración natural en sitio específica. Virus AAV

Hasta la fecha los únicos virus que de forma natural se integran en sitio específico son los virus adeno-asociados. Se trata de uno de los virus más pequeños descritos, con un genoma de 4.7Kb de tamaño, perteneciente a la familia *Parvoviridae*; dentro de los cuales están englobados en los Dependovirus, debido a que requieren de la infección de otros virus para poder completar su ciclo viral [50]. Aproximadamente el 80% de los humanos son seropositivos para AAV2 [51] si bien la infección por este tipo de virus no se ha visto asociado a ninguna patología humana conocida, incluso a pesar de que la expresión de la proteína viral rep78/68 induce parada del ciclo celular [52] y apoptosis [53]. Curiosamente, los niveles existentes de rep son los que determinan el paso entre fase lítica y latente, debido a su doble acción de regulador de la transcripción y mediador de la integración[54]. Los genomas virales permanecen en las células infectadas de manera episomal pudiendo formar concatémeros; y aunque infrecuentes, se han descrito integraciones en el genoma celular [55]. En los casos en los que se ha podido caracterizar su integración (relativa a AAV de tipo 2), se ha visto que poseen preferencia por el cromosoma humano 19 concretamente, por el locus denominado *AAVS1* (19q13.42) [56].

Su empleo como vector recombinante tiene en contra su proceso de producción y purificación, que es tedioso y complejo; y que la capacidad de portar un transgén es muy limitada, dado el pequeño tamaño de su genoma (4.7Kb totales). Aún así se trata de un virus que tiene un gran potencial como vector recombinante para su uso en terapia génica, gracias a su ausencia de integración, a los niveles mantenidos de expresión *in vivo* y al hecho de que son capaces de ser serotipados con distintas proteínas, lo que les confieren especificidad por distintos tejidos [57].

Para la integración específica de estos virus se han caracterizado una serie de elementos necesarios, como son:

- A. la proteína rep78/68 [58],
- B. las repeticiones terminales invertidas (ITRs) [59],
- C. el sitio de homología en el locus AAVS1, compuesto de un sitio de unión de rep (RBS) asociado al *terminal resolution site* (trs) [60], [61] y
- D. un elemento mínimo p5IEE, no sólo necesario, sino suficiente para una integración específica y eficiente [54].

Dichas interacciones han sido demostradas *in vitro* tanto con rep68 [62] como con rep78 [63]. También se ha visto que dichas integraciones ocurren en distancias variables del sitio RBS dentro del *locus* [64], [65], y que promueven reordenamientos en las zonas de unión existentes entre el *locus* y el genoma viral [66], [67].

## 3.2. Integración específica en sitio aceptor artificial.

### 3.2.1. Nucleasas<sup>3</sup>

La integración de un DNA exógeno en el genoma puede realizarse por varios métodos, tanto virales, como no virales según hemos visto. La problemática de este aspecto reside en el hecho concerniente al sitio de integración y los efectos adversos que ello puede provocar en el genoma residente [68], [69]. Como ya se ha comentado, la recombinación homóloga permite la edición génica aunque es un proceso que lleva asociados problemas de aplicación generalizada.

Una de las alternativas que se ha propuesto es el empleo de enzimas que promueven la rotura del DNA celular en las dos hebras (*double strand break*, DSB), en las proximidades de una diana genómica, con lo cual pueden incrementar la frecuencia de inserción o modificación génica entre 100 y 10.000 veces [70], [71] con respecto a los procesos de recombinación homóloga. Para ello se han empleado,

- (i) Enzimas que de manera natural no cortan en el genoma humano, como es el caso de **I-Sce** [72], que se ha utilizado en un emplazamiento genómico generado artificialmente[73]. Con dicho método se han llegado a alcanzar eficiencias próximas al 12%; pero arrastran el problema de la entrega y generación de un sitio artificial en el genoma.
- (ii) Enzimas quiméricas como las **nucleasas de dedos de zinc** (Zinc Finger Nucleases) [74], [75], o como las **meganucleasas** [76], [77]. Ambas han demostrado una alta eficiencia para la modificación de los *loci* diana contra los que fueron diseñadas. En concreto, se demostró una eficiencia de hasta el 50% para las ZFN [78], y de un 45% para las meganucleasas [79]. En este caso, el inconveniente reside en su escaso desarrollo, su elevado coste, la forma de entrega en la célula diana y los posibles efectos *off-target* que tienen en el genoma celular[80].
- (iii) El uso de **transposasas** [81], que sin arrastrar secuencia viral alguna, son capaces de promover la inserción de un transgén en el genoma (al azar), pero de manera reversible [82], [83]; es decir, una vez ejercido el efecto terapéutico, se podría escindir la casete, sin dejar marca génica alguna. La salvedad reside en el efecto que puedan ejercer en el sitio de integración (aunque sea de manera temporal) y la entrega, tanto de la enzima, como de la casete a integrar[84].
- (iv) Una última alternativa reside en las **recombinasas**. En este supuesto, se trata de enzimas específicas de secuencia, las cuales catalizan una reacción de reordenamiento de segmentos de DNA mediante su unión a secuencias específicas de pequeño tamaño, donde cortan el DNA y promueven el intercambio entre las distintas hebras involucradas y su unión final [85]. Una característica primordial de

---

<sup>3</sup> En realidad, dentro de las enzimas a tratar existen tanto sitio específicas natural como artificiales, así como pseudo-aleatorias, como es el caso de las transposasas. Pero se ha decidido su descripción dentro de este epígrafe dado que aporta una visión más clara que su separación y enunciado bajo distintos epígrafes.



estos sistemas es su alta especificidad, eficiencia y rapidez[86]. (Como entran en el objeto de desarrollo de esta tesis se especificarán a continuación).

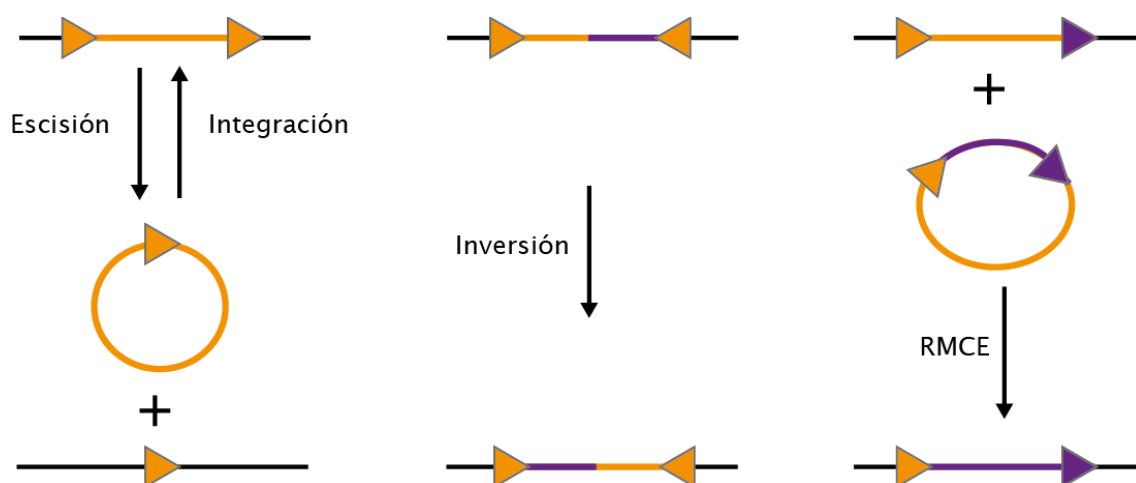
### **3.2.2. Recombinasas. Sistema de recombinación mediado por intercambio de casetes (RMCE)**

Las recombinasas son de distintos tipos, dependiendo del núcleo catalítico que posean (en concreto dependiendo del aminoácido central de este, según sea tirosina o serina), o de su origen (fagos bacterianos o levaduras). De entre todas ellas, las más caracterizadas son, PhiC31 [87] y Cre [88], derivadas de fagos bacterianos; y Flp [89], con origen en *S. cerevisiae* (levadura). Cada una de ellas reconoce distintas secuencias o sitios de recombinación (att, loxP o FRT, respectivamente), donde catalizan las distintas reacciones de recombinación [90]. Estas secuencias se componen de un núcleo central o secuencia espaciadora y dos secuencias palindrómicas que la flanquean [91].

Dependiendo de la naturaleza y la posición de dichas secuencias de recombinación, se dará lugar a alguno de los procesos que se describen a continuación (figura 2):

- A. Integración. La recombinasa se encarga de catalizar la reacción entre un donador circular (por regla general) en una localización genómica que contiene una única secuencia de reconocimiento.
- B. Resolución o escisión. Se cataliza entre dos sitios de reconocimiento emplazados en la misma hebra de DNA y con la misma orientación.
- C. Inversión. Cambia la dirección de un determinado fragmento de DNA, entre dos sitios de recombinación que se encuentran en la misma hebra, pero con orientaciones invertidas entre ellos.
- D. RMCE. Se puede considerar como la combinación de las reacciones de integración y escisión entre dos secuencias de DNA, las cuales poseen dos pares de secuencias de recombinación en la misma hebra. Como producto final se producirá el intercambio de las secuencias contenidas entre los sitios de recombinación.

Mientras que la reacción de escisión es irreversible (debido a la pérdida del producto circular escindido); tanto la inserción, como la inversión son reacciones reversibles, debido a la permanencia de los sitios de recombinación inalterados y contenidos dentro de la misma hebra de DNA. Con el fin de poder modificar un locus previamente marcado, se diseñaron estrategias que fueran irreversibles o que promovieran la reacción en una determinada dirección. Para ello se ha aplicado la tolerabilidad que estas enzimas poseen a ciertas variaciones en la región de reconocimiento, para así alterar su patrón de recombinación o prevenir la reversibilidad [92]. Dichas variaciones pueden ser de dos tipos (tabla 2), variaciones en la secuencia palindrómica, las cuales en su reacción generaran una secuencia original (o *wild type*) y una secuencia inerte [93] y por otro lado, variaciones en la secuencia espaciadora, las cuales permiten un número ilimitado de recombinaciones. Estas últimas se componen de sustituciones de nucleótidos en la región espaciadora, de tal manera que se promueve la reacción entre sitios idénticos (homotípicos) y está claramente desfavorecida, sino abolida entre sitios distintos (heterotípicos) [90]. Dicho planteamiento es la base de la reacción de intercambio de casetes mediada por la recombinasa (RMCE) [92].



**Figura 2. Posibles reacciones catalizadas por recombinasas.** Se componen de cuatro posibilidades, excisión/integración, inversión e intercambio de casetes. Los sitios de reconocimiento por recombinasa se han representado como triángulos naranjas (y morado para RMCE).

En dichas reacciones de RMCE, un marcador de selección (positivo, negativo o ambos) flanqueado por sitios heteroespecíficos es insertado en una localización cromosómica, para posteriormente ser reemplazado mediante el intercambio con un fragmento de DNA que contiene dichos sitios flanqueando un marcador o gen selectivo distinto del inicial. Para asegurar un éxito en dicha reacción los sitios heteroespecíficos no han de poder recombinar entre sí, este es el caso del lox2272 [94] y no el caso del lox511 el cual aunque baja posee una tasa de recombinación de entre el 1-5% con el sitio loxP original [95]; lo que conlleva una reacción de RMCE menos eficiente.

```

5'  ATAACTTCGTATA ATGTATGC TATACGAAGTTAT  3'
3'  TATTGAAGCATAT TACATACG ATATGCTTCAATA  5'

```

| Nombre   | Palíndromo 5' | Sec. espaciadora | Palíndromo 3' |  |
|----------|---------------|------------------|---------------|--|
| lox511   | ATAACTTGGTATA | ATGTATaC         | TATACGAAGTTAT | Variaciones en la secuencia espaciadora  |
| lox5171  | ATAACTTGGTATA | ATGTgTaC         | TATACGAAGTTAT |  |
| lox2272  | ATAACTTGGTATA | AaGTATcC         | TATACGAAGTTAT |  |
| m2       | ATAACTTGGTATA | AgaaAcca         | TATACGAAGTTAT |  |
| lox71/LE | taccgTTGGTATA | ATGTATGC         | TATACGAAGTTAT | Variaciones en la secuencia palindrómica |
| lox66/RE | ATAACTTGGTATA | ATGTATGC         | TATACGAAcggtA |  |

**Tabla 2. Variaciones de las secuencias lox reconocidas por la recombinasa cre.** Se muestran variaciones tanto en la región espaciadora como en la secuencia palindrómica. Los cambios con respecto a la secuencia loxP se muestran en caracteres rojos. Se denominan secuencias heteroespecíficas.

De entre todas las enzimas mencionadas, la más utilizada es la enzima cre, dada su funcionalidad y eficiencia en sistemas mamíferos y el amplio abanico de secuencias heteroespecíficas disponibles [90] (tabla 2); lo cual la convierte en una herramienta ideal para su utilización en sistemas que medien el intercambio de forma segura, específica e inequívoca [90].

### 3.3. Integración al azar en sitio aceptor natural. Lentivirus

#### 3.1.1 Vectores derivados de *Retroviridae*

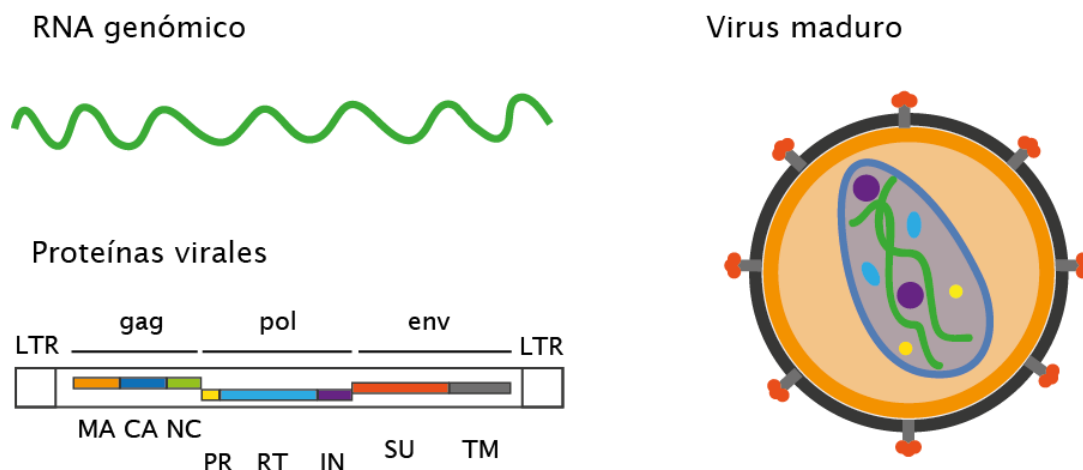
Los retrovirus se componen de un genoma diploide de RNA, el cual es de polaridad positiva y dependiente de la retro-transcriptasa, lo cual implica que el ciclo viral está supeditado a la copia del RNA viral a cDNA, para de esta manera poderse integrar (provirus) en el genoma de la célula infectada. Las células que portan dichas integraciones, lo harán de manera estable, creando por tanto una progenie celular modificada permanentemente.

Dicha propiedad de integración los convirtió en los virus diana inicial para el desarrollo de vectores virales recombinantes: el primer vector viral en ser utilizado fue el Gammaretrovirus de la leucemia murina (MLV), tanto en protocolos de terapia génica en humanos [96], como en experimentos de marcaje genético [97]. La limitación en cuanto a su uso reside en la ausencia de infectividad (y por tanto de transferencia génica) de células quiescentes [98] y el riesgo de tumorigénesis insercional [99], [100]. Como alternativa a ellos y solventando parte de estas restricciones, se han desarrollado nuevos vectores retrovirales basados en lentivirus<sup>4</sup>, principalmente en el HIV-1, dado que poseen la habilidad de infectar (o transducir) células que no se encuentran en proceso de división, y evitan posibles efectos entorno a su integración en el genoma celular, gracias a las modificaciones que se han implementado [102].

Todos los retrovirus, se organizan en tres regiones codificantes principales que regulan la expresión de las proteínas del virión, estas son: *gag*, la cual codifica para las proteínas estructurales de la matriz, cápsida y nucleoproteínas; *pol*, que a su vez codifica para la retro-transcriptasa, la integrasa y la RNasaH; y *env*, de la que se derivan los componentes transmembrana y de superficie de la envuelta proteica viral [103]; además poseen una proteasa viral, que dependiendo del tipo de retrovirus se encuentra emplazada en *gag* o *pol* (figura 3). Los lentivirus codifican para una serie de proteínas pequeñas accesorias[104]: *vif*, *nef*, *env*, *vpu*, *vpr*, *tat* y *rev*; expresadas por *splicing* alternativo; éstas están implicadas en procesos de entrada y salida, regulación génica y control de actividades de proteínas celulares, lo que confiere una mayor complejidad en el ciclo viral *in vivo* e *in vitro*.

---

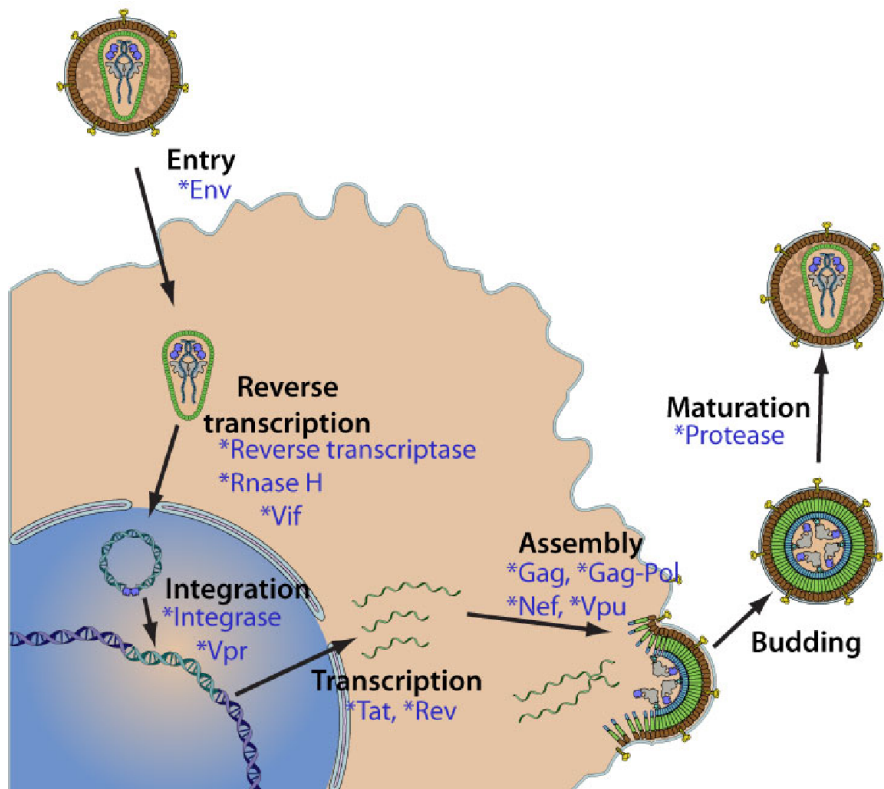
<sup>4</sup> Inicialmente fueron desarrollados para la utilización en células del sistema nervioso central, dada su capacidad de transducir células que no se encontraban en período de división [101], aportando ventaja sobre los sistemas retrovirales.



**Figura 3. Componentes principales de los retrovirus.** a) RNA genómico viral, b) distintas proteínas codificadas y necesarias para la formación del virión y c) conformación del retrovirus maduro.

Las partículas virales son de forma esférica, externamente constituida por una envuelta lipídica procedente de la membrana plasmática celular; en ella, se hayan ancladas las glicoproteínas env. Hacia el interior de la partícula se localiza la matriz formada por la proteína p17 (MA), que interacciona doblemente con la envuelta y con la cápside, que posee una morfología cónica formada por p24 (CA). En ella, se encierran las dos copias de RNA genómico, las cuales interaccionan cada 5-6 nucleótidos con una molécula de la proteína básica de la nucleocápsida p7 (NC). El virión contiene en su interior alrededor de 750 copias de proteínas implicadas en la replicación del genoma viral (pol) y al menos 5000 moléculas de gag [105].

El dominio gag, que se sintetiza como una poliproteína (MA-CA-NC) refleja la organización física en los viriones, estos complejos altamente ordenados en trímeros de MA y dímeros de CA, hacen de chaperón para la interacción y segregación de las dos copias de RNA genómico (RNAg). Dichos complejos (proteína-RNAg) quedan anclados por un residuo miristoilado de MA con la membrana plasmática en *lipid-rafts* ricos en env. El *budding* activa la proteasa viral pro que procesa, ya en la partícula liberada, las poliproteínas gag-pol, que se reordenan en cápsida y nucleocápsida mediante auto-ensamblaje de las subunidades procesadas de p17, p24, p7; dando lugar a la cápsida madura con forma de bala (o cónica) [104] (figura 4). El ciclo viral, implica la entrada por fusión de membrana mediada por receptor, lo que libera la cápsida en el citoplasma celular que parcialmente desensamblada constituye el complejo de pre-integración (PIC) donde tiene lugar la retrotranscripción mediada por las proteínas aportadas en el virión. El cDNA sintetizado es transportado en el PIC al núcleo celular mediado por la secuencia de translocación nuclear contenidas al menos en la integrasa, y otras proteínas accesorias (vpr) y por la secuencia en cis en el cDNA (*flap*). El provirus integrado por la acción de la integrasa se compondrá de las LTR (*Long Terminal Repeats*), que se encuentran en los extremos del sitio de integración, englobando los genes virales (*gag-pol* y *env*); los cuales transcribirá por la actividad promotora que posee [106] (figura 4).



**Figura 4. Ciclo viral de HIV-1.** En el se representan las distintas fases en el proceso de infección lentiviral hasta la producción de nuevas partículas virales desde la célula infectada. Imagen tomada de <http://www.viralzone.expasy.org/>.

Entre las ventajas que presentan como vectores de transferencia destacan:

- A. capacidad de carga (hasta 10 kpb)
- B. facilidad de manipulación y producción
- C. disponibilidad y posibilidad de pseudotipado<sup>5</sup>,
- D. período prolongado en la expresión del transgén, derivado del hecho de su permanente integración.
- E. tolerancia en sistemas de expresión múltiple (policistrones, promotores, IRES)
- F. toleran empaquetar RNAs con una corta secuencia de empaquetamiento
- G. contienen las enzimas requeridas para la fase inicial de su ciclo incorporado en el virón.

Su desarrollo e implantación como herramienta de uso cotidiano en los laboratorios de biología molecular se debió a la posibilidad de generar vectores retrovirales no competentes; gracias al desarrollo de líneas empaquetadoras, que aportaban las proteínas retrovirales necesarias en *trans* [109], [110]. Unido al hecho de que las secuencias que

<sup>5</sup> Como se mencionó con anterioridad, el pseudotipado de los vectores con proteínas exógenas [107], supone un avance en cuanto a su aplicabilidad. En el caso de los vectores retrovirales el uso de la proteína VSV-G (*vesicular stomatitis virus glycoprotein*), ha permitido generar vectores más estables, permitiendo una concentración eficiente de las partículas virales y un amplio tropismo [101], [108].

codifican para dichas proteínas dejan cabida para la introducción de secuencias exógenas [111]. Esto ha sido posible debido al hecho de que los primeros pasos tras la entrada en la célula diana, retrotranscripción e integración, no necesitan de síntesis de proteínas *de novo*, lo cual puede ser aprovechado para una entrega eficiente de proteínas de interés o la visualización del proceso de transducción[112].

Además, si son comparados con otros vectores virales presentan ventajas significativas[113], [114]:

- A. poseen la habilidad de transducir un amplio espectro de tipos celulares e incluso gran variedad de especies animales,
- B. son capaces de integrar el material génico portado en las células diana,
- C. pueden modular los niveles de expresión del transgén en niveles dependiendo del promotor utilizado,
- D. la probabilidad de generar partículas competentes, derivado de la recombinación entre los distintos plásmidos es mínima, y
- E. presentan una facilidad de producción y escalado.

### **3.1.2 Sistemas de generación de partículas lentivirales recombinantes**

Se describe en detalle en la sección de Materiales y Métodos.

### **3.1.3 La integración lentiviral**

Tras la entrada del RNA viral en la célula, este es retrotranscrito en el citoplasma mediante el complejo de pre-integración (PIC). Dicho complejo es capaz de atravesar la membrana nuclear [115] y una vez allí, llevar a cabo el proceso de integración (entre el vector cDNA y el DNA celular) mediado por la integrasa del provirus (figura 4).

La integrasa (*in*) es una proteína de 32 KDa, codificada como parte del complejo multiproteico *gag-pol* y escindida por la acción de la proteasa [116] en el proceso de maduración del virión. Aunque no específica, su acción se ha descrito que no es totalmente aleatoria; sino que hay determinadas zonas preferidas sobre otras, variando la tasa de integración dos órdenes de magnitud entre las primeras (o *hotspots*) y las integraciones aleatorias [42].

A pesar de la eficacia de la reacción de integración [116] una fracción del cDNA viral, no se integra [117], [118]; sino que una vez internalizado en el núcleo puede permanecer en estado episomal en formato 1 LTR, que se generará por la recombinación homóloga entre ambas LTRs [119] o en forma de episoma circular de 2 LTRs, por el proceso de unión de extremos no homólogos (*NonHomologous End Joining*, NHEJ) [120]. Además, un tercer intermedio generado es el cDNA de doble hebra en formato lineal [121].

La cantidad de estos intermedios que existen en células en división se ha cuantificado ampliamente y por diversos métodos [122]; de manera general se ha observado que la ratio entre las tres formas posibles (las dos circulares anteriormente mencionadas y la forma lineal integrada, pues la lineal no integrada es rápidamente degradada por la maquinaria celular) transcurridas 48 horas tras la infección es aproximadamente de 55% lineal

integrado, 35% circular 1LTR y una fracción inferior al 10% se compone de circular de 2LTRs [123].

### 3.4. Modificación no-integrativa limitada<sup>7</sup>

#### 3.4.1. Vectores lentivirales integrasa deficientes (IDLVs)

La inserción génica retroviral puede desencadenar problemas en el genoma residente, como son la mutagénesis insercional [124], la activación de genes celulares por inserción del promotor [125] e incluso su posible efecto deletéreo dada la entrega de genes con carácter tóxico [100]. Por ello se han estudiado las distintas formas de generar vectores lentivirales deficientes en integración, lo que permitiría beneficiarse del potencial episomal de los intermedios 1-LTR o 2-LTR (anteriormente mencionados) y sus aplicaciones [126-128], dado que retienen intactas todas las capacidades lentivirales, a excepción de la integración.

Se han analizado distintas vías para tratar de evitar la integración lentiviral: la integrasa tiene un papel proactivo en la transcripción reversa y en el transporte del PIC al núcleo de la célula [129], es por ello que no puede ser eliminada en el proceso de generación de lentivirus recombinantes. Sin embargo, se ha visto que la introducción de mutaciones puntuales en la enzima [130], si son toleradas tanto en el proceso de generación de virus recombinantes, como posteriormente en la infección y expresión viral; teniendo efectos adversos mínimos en la obtención de virus recombinantes [131].

Las mutaciones en dicho gen son clasificadas en dos grupos, (i) las mutaciones de tipo I, que afectan a la integración del vector directamente y (ii) las mutaciones de tipo II, que tienen efectos deletéreos en el empaquetamiento de las partículas virales o en el proceso de retrotranscripción o en la translocación del PIC; derivando finalmente en una ausencia de integración [130]. Es por ello, que para favorecer el establecimiento episomal de los intermedios lentivirales conviene utilizar mutaciones que no afecten al resto de sus propiedades, sino únicamente a su capacidad de integración; por tanto restringiendo su uso a las de tipo I, para no afectar a la obtención y posterior transducción con las partículas recombinantes virales [132].

La integrasa se compone de tres dominios funcionales: el extremo *N terminal* o dedo de zinc (*Zinc Finger Domain*), el núcleo catalítico y el extremo *C terminal* o de unión a DNA [133]. El núcleo catalítico contiene una secuencia aminoacídica que es necesaria para la correcta actuación en el proceso de integración [134]: se trata del motivo D-D-E, que se compone del aspártico D64, aspártico D116 y glutámico E152. Estos aminoácidos, no sólo están conservados por las integrasas de la familia *Retroviridae*, sino que también se encuentran en las transposasas bacterianas[135], sugiriendo la importancia funcional del motivo.

---

<sup>7</sup> No se considera la expresión mediada por formas episomales generadas mediante el empleo de vectores derivados de Adenovirus o AAV ya que están fuera del objeto de estudio del trabajo de tesis llevado a cabo.

Distintas mutaciones en cada uno de los dominios funcionales, pueden ser llevadas a cabo con la finalidad de generar vectores lentivirales integrasa defectivos o deficientes (IDLVs) [136]; pero como se ha comentado únicamente las de tipo I son útiles para la generación de vectores virales funcionales. Por tanto, las más desarrolladas, son las que afectan a la tríada presente en el núcleo catalítico y en concreto, D64 y D116 puesto que inactivan las propiedades integrativas de la enzima. En concreto la modificación de aspártico por valina (D64V) y de aspártico por isoleucina (D116I), son las más empleadas, puesto que no suponen una reducción en la síntesis de cDNA y por tanto resultan idóneas para su empleo como IDLVs[132].

La tasa de permanencia en la célula transducida por parte de los IDLVs, es limitada e inversamente proporcional a la de división celular (debido a la ausencia de un origen de replicación). Sin embargo, se han mostrado eficaces, en células quiescentes o con baja tasa de replicación ya que pueden permanecer durante el tiempo necesario para la actuación requerida [103], [137].

Por tanto y dado su grado de desarrollo se dispone de la posibilidad de utilizar un sistema lentiviral de entrega de información genética, pero solventado el problema asociado a su naturaleza integrativa.

### **3.5. Modificación no-integrativa no-limitada**

#### **3.5.1 Episomas replicativos**

Desde finales de los años 90 se cuenta con la tecnología de los cromosomas artificiales humanos (HAC) [138], que representan la mejor manera de mimetizar la expresión génica desde un entorno genómico natural, es decir, cromosómico. Sin embargo, las dificultades, tanto en su manipulación, como en su entrega, son evidentes debido a su gran tamaño (entre 500 Kb y 2 Mb) [139]. Por esta dificultad se hace necesario el desarrollo de nuevos vectores, más versátiles, que posean una entrega en las células diana eficiente y sostenida; pero con la expresión génica regulada desde un sistema extra-cromosómico [140].

Los episomas replicativos se han postulado como una alternativa segura para el empleo en transferencia génica sin interferencia, al permanecer como DNA extra-cromosómico. La replicación de estas formas episomales permite su mantenimiento “ilimitado” al contener orígenes de replicación (ori). Los primeros orígenes caracterizados tienen una procedencia viral; se trata de formas episomales presentes en el ciclo de pequeños virus con genoma de DNA (*Polyomaviridae*, *Papillomaviridae*). Además se han descrito formas latentes de algunos *Herpesviridae* (Epstein-Barr virus), que también poseen un ciclo episomal. El estudio de todos ellos, ha permitido caracterizar *Oris* y proteínas virales implicadas en su mantenimiento (*large T-antigen* o E1/E2 y EBNA por citar un ejemplo de cada grupo mencionado). Derivado de esta circunstancia, el empleo de episomas en modificación génica se propuso principalmente en aplicaciones de terapia oncolítica viral, es decir, el tratamiento de células tumorales mediante el empleo de vectores virales replicativos



condicionales [141]. El riesgo de transformación asociado al potencial oncogénico de las proteínas virales necesarias en la replicación viral [142], [143] y el uso limitado que se les ha atribuido durante años, han provocado que su utilización en aplicaciones de transferencia génica se haya visto desarrollada en menor medida.

Gracias a los avances realizados en las últimas décadas, se ha establecido que un episoma replicativo ha de ser capaz de replicarse una vez por ciclo celular [144] y deberá contener (además de los genes exógenos que se pretenda entregar), distintos elementos reguladores, que ayuden

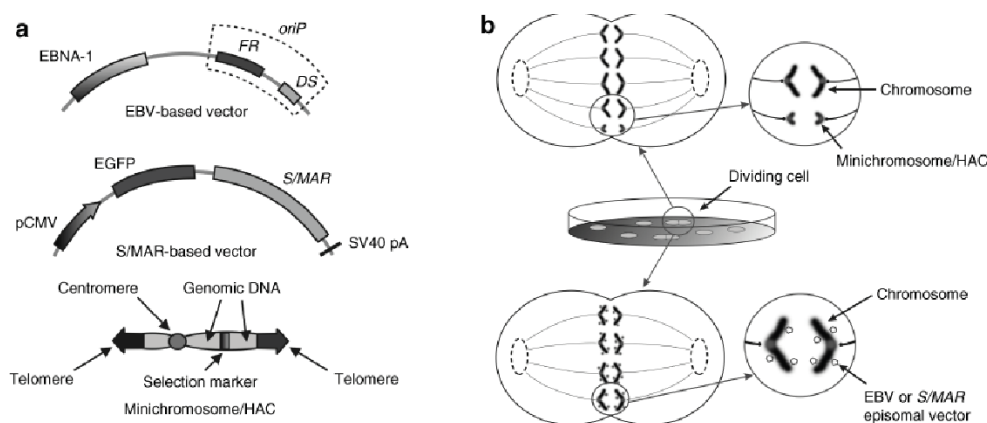
- A. tanto al mantenimiento de los vectores, entre los que se incluyen los elementos de unión a la matriz o andamiaje nuclear (S/MAR) y los elementos de la región del control del locus (*Locus Control Región*, LCR) [145],
- B. como secuencias que sirvan de origen de replicación [146] y
- C. adicionalmente elementos que regulen la segregación episomal [147].

El plásmido o vector ha de ser mantenido en cultivos celulares (1) por un número prolongado de generaciones y en ausencia de presión selectiva [148]; (2) deberá mantenerse en un bajo número de copias por célula (menor a 10), replicándose y segregándose eficientemente [149], [150]; y (3) se deberá evitar el empleo de elementos sujetos a regulación no deseada, como eventos de silenciamiento o metilación, lo cual se consigue minimizando los componentes bacterianos y virales [151]. Todo ello conducirá a una óptima regulación de la expresión génica desde el episoma [152].

El objetivo obliga a aunar en un plásmido o vector aquellos elementos que pudieran solventar los problemas asociados a la funcionalidad que cada uno de estos componentes tiene por separado, entre los que estarían:

- A. la baja tasa de establecimiento episomal y el problema de entrega en determinados tipos celulares o determinados tejidos (in vivo), como es el caso de los plásmidos no virales con secuencias S/MAR [148].
- B. el problema de entrega y manejo de los constructos, crítico tanto en los sistemas de amplicones de herpesvirus [153] y adenovirus [154], como en el caso de los cromosomas artificiales humanos [155].
- C. tratar de evitar la formación y mantenimiento de virus competentemente replicativos [156].

Por tanto el vector episomal deberá ser plenamente funcional en células de mamífero, deberá estar compuesto en su mayoría por secuencias de DNA eucariotas y poseer un sistema de entrega eficaz y universal (principal inconveniente que presentan los elementos episomales no-virales). Todos estos elementos podrán contribuir a una transferencia génica más eficiente y segura.



**Figura 5. Vectores episomales y su mecanismo de mantenimiento.** (A) Distintos tipos de vectores episomales utilizados y sus componentes estructurales básicos. (B) Distintos mecanismos de mantenimiento, bien sea por segregación centromérica (HAC) o por asociación con la cromatina (EBV y S/MAR). Imagen tomada de *Advances in high-capacity extrachromosomal vector technology: episomal maintenance, vector delivery and transgene expression* [139].

### 3.6. ¿Necesitamos siempre transferir genes?. Sistemas de entrega de proteínas

La entrega de transgenes persiguen en multitud de procedimientos la expresión transitoria del mismo. Por tanto, se estima la posibilidad de transferir directamente el producto proteico codificado, en lugar del material génico codificante. De esta manera se han desarrollado métodos no genéticos de entrega de liposomas y polímeros [157] y la utilización de péptidos transmembrana de fusión, que son capaces de entregar la proteína en el interior celular [158]. Este último proceso, que se denomina “transducción proteica”, puede llevarse a cabo por medio de la fusión de dominios tanto virales [159], como celulares [160]. El inconveniente reside en los sistemas de producción, por lo general bacterianos; y en el posterior aislamiento y purificación, que conllevan un inadecuado procesamiento (modificaciones postraduccionales ausentes en sistemas bacterianos) o su posterior bloqueo en los endosomas[161].

Una posible alternativa a ello es el empleo de sistemas virales capaces de englobar proteínas en sus viriones. En este sentido, los vectores lentivirales se postulan una vez más, como candidatos con eficacia probada (transducción proteica mediada por lentivirus). El primero de los sistemas desarrollados se basó en el aporte de la proteína en *trans*, en el sistema de empaquetamiento [162], [163], aprovechando la interacción que establece la proteína accesoria de HIV *vpr* (implicada en *budding* de viriones), con el dominio p6 de la proteína *gag* para su empaquetamiento [164]. La ventaja que aporta este sistema, es la ausencia de modificación en los genes estructurales de los vectores lentivirales disponibles; sino que pasa por generar plásmidos con fusión *vpr*-proteína de interés, para su expresión en el sistema de producción y así ser encapsidada. Esta estrategia se ha utilizado con éxito en el empaquetamiento y entrega de distintas proteínas, desde marcadores fluorescentes [112], hasta proteínas con aplicación terapéutica[163].

Dicho sistema presenta como principal inconveniente (además de dificultades que puede conllevar el añadir un plásmido adicional en el proceso de producción) los efectos deletéreos asociados al uso de vpr como proteína de entrega, pues se ha descrito su implicación en procesos de parada de ciclo celular [165] y citotoxicidad [166]; lo que puede traducirse en una disminución en la viabilidad de las células productoras y en la generación deficiente de partículas virales. De igual manera, la presencia de vpr en la célula diana, puede provocar efectos no deseados derivados de dicha toxicidad. De hecho como se comenta en Materiales y Métodos (Sistema de producción de lentivirus recombinantes), los vectores de tercera generación han eliminado entre otros el gen que codifica para dicha proteína [167].

La síntesis de las proteínas estructurales gag como una poliproteína con la serie MA(p17)-CA(p24)-NC(p7) (matriz-cápsida-nucleocápsida) como se mencionó anteriormente, ha permitido explorar la entrega de proteínas exógenas fusionadas en esta poliproteína. La morfogénesis viral tiene lugar en viriones libres tras la activación de la proteasa (gen *pol*) encapsidada. Cuantitativamente el sistema es de alto rendimiento: de manera natural se produce gag y gag-pol en una ratio 20:1, desde un mismo cDNA, gracias a la señal de salto de marco de lectura, que se encuentra codificada entre ambos genes [168]. Esta ratio permite el empaquetamiento de un número elevado de proteínas -aproximadamente 5000 proteínas (*gag*) o 250 partículas proteicas (*pol*) [105]-. Ello posibilita el disponer de un sistema flexible de empaquetamiento de proteínas transductoras, aunque su modificación puede dar lugar a una reducción en la productividad e infectividad viral [169].

Combinado con la posibilidad que ofrecen los virus de modular el número de partículas entregadas por célula conforme a la MOI [169]. La posibilidad de incluir sitios adicionales de procesamiento en la poliproteína, abre la posibilidad de generar proteínas recombinantes “nativas” (no fusionadas), si se sitúan flanqueando el ORF las dianas de proteasa viral.

## 4. Limitaciones y Problemas

Se han revisado aspectos de la entrega de DNA exógeno, y las alternativas de las que tecnológicamente se disponen [170]. La modificación génica puede conducir al menos a dos grandes grupos de fenómenos que han de contemplarse tanto desde la perspectiva de la bioseguridad y su aplicación clínica, como del modelo experimental en sí, que permite un análisis certero del propio ensayo. Una limitación fundamental para el desarrollo de procedimientos y modelos de transferencia génica con validez translacional, radica en la incapacidad para realizar experimentos bajo condiciones genéticas definidas[23].

La introducción de un transgén generalmente resulta en la alteración del número de copias génicas celulares, con efectos que están poco estudiados y por tanto pobremente comprendidos. Los análisis mas exhaustivos se han realizado in vitro en levaduras [171] y en *Drosophila* [172]. En relación con la expresión génica en mamíferos, datos de

transcriptómica sugieren que el número de transcritos de un gen incrementa groseramente en proporción al número de copias [173]. Es por tanto posible, que las aproximaciones de transgénesis puedan perturbar el sistema biológico, debido a que promueven un elevado número de copias en la célula. Merece mención un estudio que mediante una aproximación matemática sugiere que el incremento en la dosis génica altera la dinámica de la expresión génica [174], lo que sugiere que el efecto del número de copias sea específico del gen y del sistema. Por tanto, queda estudiar sistemáticamente el efecto de la dosis génica del transgén, que permanece desconocido en la mayoría de los sistemas, pues con el número de copias introducidas, el sistema de expresión variará.

La otra gran complejidad puede derivarse de efectos debidos de la integración. La localización en el genoma del gen ectópico puede tener profundas consecuencias en la expresión génica, tanto del transgén en sí, como de las regiones genómicas que lo rodean. La integración en zonas de heterocromatina compactada, suele dar como resultado una expresión baja o nula [175]. También conlleva el silenciamiento gradual del transgén concomitante con la hetero-cromatinización, siendo extensible a la mayoría de las copias integradas [176], especialmente si éste contiene secuencia bacterianas [177]. La integración puede por último, alterar secuencias endógenas incluyendo genes y secuencias reguladoras, generando efectos impredecibles y cambios fenotípicos no previstos [178]. Debido a estos efectos posicionales es necesario generar y analizar varias (impredecible en número) líneas transgénicas.

Ambos problemas son comunes en el empleo de los sistemas virales integrativos; por ejemplo, el que la carga viral empleada condiciona el número de copias integradas en una única célula, se ha demostrado que promueve entre otros, un cambio en la capacidad diferenciadora de células pluripotentes múridas, probablemente debido al daño provocado en el DNA [179]. Aunque estos efectos pueden resolverse modulando finamente la transducción y titulando adecuadamente los virus, la eficiencia de transducción específica condiciona en ocasiones el empleo de cargas virales elevadas que favorecen dichos procesos. La variación clonal en la expresión del transgén, y la ruptura génica pueden resolverse mediante el uso de recombinasas y RMCE para introducir el transgén en un loci genéticamente “estable” y transcripcionalmente activo, este es el caso del locus ROSA26 múrido [180].

El desarrollo de vectores con características adecuadas que solventen estos inconvenientes se ha llevado a cabo mediante la generación de vectores en los que la actividad de la integrasa esté unida a la inserción del genoma en un emplazamiento predeterminado y seguro [181]; combinando la alta transducibilidad de estos virus con elementos génicos para su mantenimiento estable a modo de DNA terapéutico [182], [183].

Focalizando la atención en los sistemas retro/lentivirales, que es uno de los sistemas más desarrollados, las estrategias que se han intentado para modular su integración en sitio específico o dirigida, se ha centrado principalmente en

- A. fusiones con la integrasa endógena viral [184-186],
- B. vectores híbridos para la entrega de enzimas de dedos de zinc (*Zinc Finger Nucleases*, ZFN) [78],
- C. sistema lentiviral híbrido con transposasas, *Sleeping Beauty* [82], [83] y
- D. sistema de inserción mediada por recombinación (RMCE), con la utilización de la enzima *Flpasa* (Flp) [187] o cre [188].

Sin embargo, todas estas aproximaciones mantienen algún tipo de dificultad técnica no resuelta en su totalidad; lo cual impide el disponer de un sistema óptimo y de uso, si no universal, amplio. En particular se requieren mejoras en cuestiones de eficiencia, seguridad y efectividad. Brevemente, sin ánimo de ser exhaustivo: las transposasas aunque escindibles tienen un alto porcentaje (30%) de inserciones en sitios o regiones transcripcionalmente activas [189]; las fusiones con las integrasas han sido incapaces de llegar a eficiencias elevadas, tanto en el empaquetamiento, como en la seguridad del locus escogido [186]; la tecnología basada en nucleasas adolece de la dificultad del diseño, el elevado coste económico y la presencia de efectos colaterales derivados de cortes inespecíficos (*off targets*) [190] o en el caso de las enzimas recombinasas, el sistema diseñado o elegido no esclarece finamente y de manera concisa los eventos de recombinación falsos positivos [91], [191].

Parece por tanto oportuno el desarrollo de nuevos sistemas en los que (i) tanto los eventos dirigidos de inserción sean sobre un emplazamiento seguro, transcripcionalmente activo y estable en multitud de estados y de tipos celulares, (ii) se permita discernir fácilmente los eventos fidedignos de los que no lo son, y (iii) desarrollar alternativas eficaces a los sistemas de integración dirigida, mediante el estudio de sistemas de mínima interferencia, episomas replicativos.

En conclusión, el estado de la cuestión del *targeting* génico específico, seguro y de un número de copias controlado (único) es uno de los campos de actuación más vibrantes de la modificación génica de eucariotas. Un territorio donde aún queda por desarrollar tecnología eficaz que permita la aplicación de protocolos de intervención génica segura y eficaz.



## **OBJETIVOS**





## Hipótesis

La combinación de vectores lentivirales deficiente en integración con

1. sistemas de recombinasas y nucleasas de zinc,
2. elementos de mantenimiento episomal, y
3. procedimientos de entrega no génica de enzimas

puede permitir el desarrollo de plataformas de elevada capacidad para promover de forma eficaz y sencilla la modificación de genomas en sitio específico y en número de copias controlado y eventualmente único.

## Objetivos

1. Generar y verificar un sistema celular versátil y sencillo para el intercambio eficaz de casetes génicos mediante la recombinasa cre utilizando vectores lentivirales
2. Desarrollar vectores lentivirales para la entrega de enzimas funcionales aplicables en modificación génica
3. Diseñar, generar y probar un modelo episomal replicativo  
para modificación celular de mínima interferencia



## **RESULTADOS**



## MODELO 1

### SISTEMA DE INTERCAMBIO DE CASETE GÉNICA MEDIADO POR RECOMBINASAS (RMCE) Y LENTIVIRUS DEFICIENTES EN INTEGRACIÓN (IDL<sub>V</sub>)

#### 1. Objetivo del diseño

Como se ha comentado en la introducción, la reacción de intercambio de casetes mediada por recombinasa (del inglés RMCE), permite la integración de un gen de interés (GOI) en una localización cromosómica específica si posee ciertas características (anteriormente detalladas), y que se traduce en ventajas sobre el empleo de técnicas convencionales:

- La expresión de un GOI en una localización determinada y pre-caracterizada aporta robustez y resta variabilidad, causada por el efecto de inserciones al azar y múltiples.
- El empleo reiterado del sitio de integración al evitar efectos posicionales aporta homogeneidad en los estudios comparativos.
- La validación y reutilización de un mismo sitio reducirá el tiempo de análisis, número de muestras y simplificará el proceso de verificación.

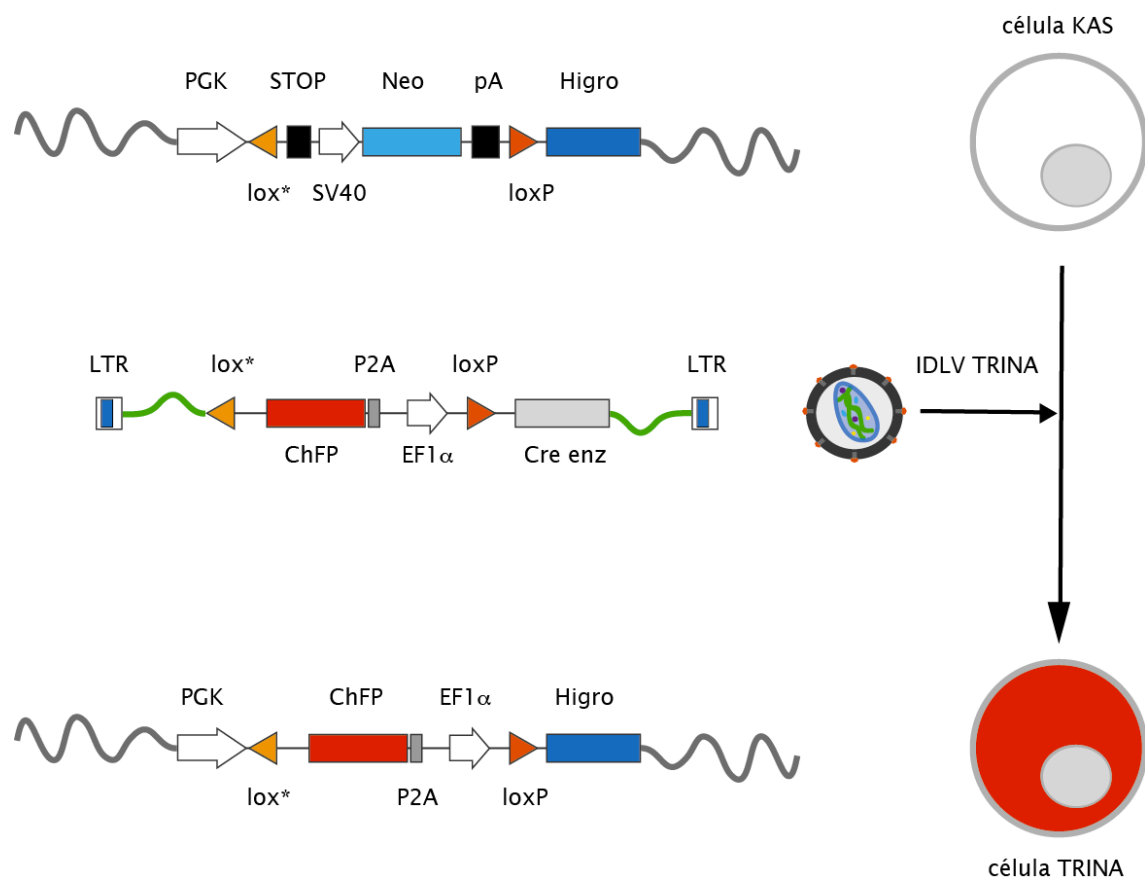
Con todo ello, los puntos que se quieren desarrollar con el presente estudio son (i) el diseño de un casete versátil y funcional tanto en la expresión génica como en la eficiencia de recombinación; y (ii) la elección y uso de un sitio génico seguro (en la manera en que pueda definirse inequívocamente) que además contribuya a favorecer (i).

El objetivo es generar clones de células que posean dichas propiedades que las hagan susceptibles a RMCE como portadoras de la casete aceptora y que serán líneas fundadoras sobre las que trabajar. Mediante la entrega de la casete donadora se permitirá (i) llevar a cabo sucesivos reemplazamientos y (ii) facilitar la detección y análisis de una manera rápida y eficaz de los eventos de recombinación fidedignos (positivos), minimizando el número de casos en los que esta recombinación fiel no se detecte o puedan quedar enmascarados por resultados negativos. Eventualmente el sistema permitiría el desarrollo de un procedimiento para generar distintas líneas celulares que sean manipulables de manera sencilla, rápida y eficaz.

#### 2. Elementos del diseño

##### 2.1. RMCE

El sistema diseñado se basa en la RMCE; atendiendo a esto, la estrategia seguida y probada en su funcionalidad en el modelo propuesto está centrada en (i) un emplazamiento génico preciso, (ii) vectores de transferencia lentivirales deficientes en integración, y (iii) secuencias de reconocimiento por recombinasa que garanticen la reacción en la forma y dirección proyectada.



**Figura 6. Sistema KAS/TRINA.** Representación esquemática de los elementos génicos contenidos en KAS (arriba) y TRINA (centro) y la estructura y composición del producto de RMCE entre ellos (abajo). Los promotores se representan por bloques con punta, los ORFs mediante rectángulos, y las secuencias lox como triángulos. El esquema de la derecha remarca la transición de célula neoR (gris) a ChFP+ (rojo) tras la transferencia de TRINA por IDLV

## 2.2. Selección de los eventos favorables

El sistema, representado en la figura 6, está basado en la estrategia tradicional de *promoter trapping*. Dicha estrategia consiste en un casete de *gene trapping* que se compone de un marcador genético (marcador de selección), seguido de una señal de poliadenilación y carente de promotor, para de esta manera utilizar el promotor endógeno propio del sitio de inserción de la casete. Se han añadido o modificado algunos de los elementos para favorecer la eficacia, fidelidad y sencillez en la generación y selección de eventos de recombinación:

- el uso de casetes intercambiables flanqueados por secuencias de recombinación heteroespecíficas (ver Introducción, apartado Recombinasas) fomentará el intercambio (reacción inter-cadenas) en lugar de la escisión (reacción intra-cadena).
- promotores exógenos sin ORF *downstream* que permitan aprovechar los eventos de recombinación no sólo para rescatar la expresión de los marcadores de selección, sino también para la expresión de distintos elementos génicos.
- la selección por un doble marcaje cuya expresión se rescata únicamente tras el evento de recombinación y que está constituido por (i) un gen de resistencia y (ii) la expresión de una proteína fluorescente que facilita el seguimiento tras la reacción.

Todo ello permitirá tener dos vías posibles para poder seleccionar las células deseadas y minimizar o eliminar el rescate de células en las que se hayan producido eventos de recombinación no fidedignos y/o al azar (falsos positivos) que enmascararían la reacción de RMCE.

### **2.3. Transferencia lentiviral**

El empleo de sistemas de vectores lentivirales permitirá una entrega de construcciones génicas de forma menos deletérea para la célula que las estrategias comunes de electroporación o nucleofección; y controlable en su temporalidad mediante la multiplicidad de infección (MOI). Su fenotipo deficiente en integración minimiza su capacidad de interferencia por inserción al azar.

### **2.4. Expresión limitada**

Como se comentó en la introducción, las recombinasas pueden inducir daños celulares promovidos por su acción sobre el genoma. Por esta razón el diseño sobre el que se ha trabajado incluyó una estrategia que pudiera moldear la expresión de la recombinasa. Para ello es la propia actividad de la enzima la que escinde la casete de expresión al estar ésta flanqueada por secuencias diana de la misma. La vida media de los productos que intervienen (molde transcripcional activo, mRNA y la proteína) serán los determinantes del periodo de actividad de la recombinasa y por tanto de su toxicidad.

### **2.5. Descripción funcional del diseño**

#### **2.5.1. Diseño del casete aceptor: KAS1.0**

Para el desarrollo de la estrategia de RMCE, en primer lugar se han de obtener células que porten los sitios de recombinación necesarios, a ambos extremos del casete que se quiere intercambiar. En dicho vector se ha introducido una casete de selección entre dos secuencias lox (“floxeado”) formada por la secuencia que codifica para la aminoglucósido fosfotransferasa (*neo*) que confiere resistencia a geneticina (G418) bajo el control del promotor de SV40 (*Simian Virus 40*) y la señal de poliadenilación de SV40. El promotor viene precedido de una potente secuencia stop para minimizar/impedir cualquier expresión proveniente de promotores situados *upstream* en el genoma. Todo ello constituye nuestro casete de intercambio y nos permitirá seleccionar las células marcadas frente a las que no lo estén, es decir, generar líneas estables sobre las que producir los intercambios.

Externo en 5' a dicha casete se ha dispuesto el promotor de la fosfoglicerinato kinasa (PGK), un promotor ubicuo de actividad media y funcional en multitud de tipos celulares [192] y para el que no se ha descrito silenciamiento en estadios de desarrollo. La secuencia stop situada en la casete *downstream* y justo 3' del loxP de la casete de intercambio (figura 6), impedirá dirigir la expresión de cualquier elemento génico celular. En el extremo 3' (externo al casete de intercambio) se ha clonado el cDNA de la fosfotransferasa de higromicina B (*hph*) que codifica para la resistencia a higromicina (hygroR), pero sin promotor, lo cual unido a la presencia de la señal de poliadenilación de SV40 en la región 5'

adyacente, otorga un marcador no funcional en la célula diana. Estas células marcadas se denominarán células KAS (*Knock-in/out Acceptor Site*).

### 2.5.2. Diseño del casete donador: TRINA1.0

El impacto sobre las células se realizará mediante los IDLV, para llevar a cabo la reacción de intercambio de casete. Esta circunstancia ha de reducir la integración al azar de los intermedios de la reacción de retrotranscripción que unido al proceso de recombinación, eliminará los efectos producidos por inserciones aleatorias en el genoma celular.

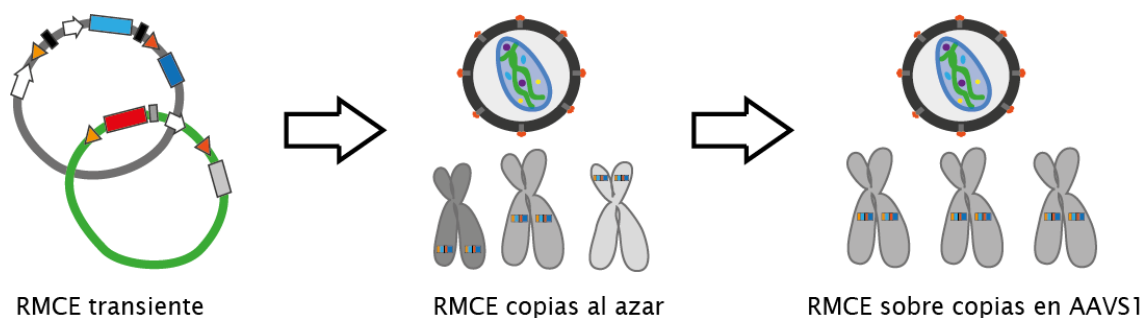
La construcción lanzadera contiene además de los elementos reguladores y esenciales para producir lentivirus (ver apartado de M&M), una casete de transferencia que en la dirección 5' -- 3' porta el cDNA que codifica para la proteína roja fluorescente (mChFP, mCherry) sin región promotora *upstream* y el promotor EF1alpha (EF1a), dichos elementos se encuentran flanqueados por los sitios heteroespecíficos lox2722 y loxP, respectivamente. Tras la secuencia promotora y el loxP se ha dispuesto el gen de la recombinasa cre. El promotor EF1a dirige la expresión de la enzima cre recombinasa y únicamente como resultado de la recombinación con la casete aceptora pasará a gobernar la expresión del gen de resistencia para higromicina, silente en las células KAS; además, dicho intercambio, interrumpirá la expresión de la recombinasa ya que el promotor será escindido de su posición. Esta reacción generará células recombinantes que serán ChFP+ (rojas) y resistentes para higromicina (higroR), perdiendo la selección para neomicina (figura 6). Estas células se denominarán *TRINA* (*Transfer into Acceptor*) al igual que los vectores utilizados para su obtención.

### 2.5.3. Generación de los elementos génicos y biológicos

Una vez diseñados los elementos necesarios para llevar a cabo el proceso de recombinación, se generará un modelo celular sobre el que trabajar que permita realizar las pruebas de función. Se ha elegido la línea celular humana HEK293A, una línea inmortalizada y de uso generalizado en el laboratorio, por su facilidad en los procesos de transfección e infección, lo cual aportará eficiencia además de una labor de mantenimiento poco laboriosa.

En segundo lugar, se valorará la prueba de función en distintas condiciones que son resultado del proceso de introducción de las casetes de intercambio (plásmidos de recombinación por transfección o infección) y en que formato intracelular permanecerán dichos elementos (plásmidos transientes, copias integradas al azar, copia en locus predefinido y en copia única). En consecuencia se ha definido un modelo de validación progresivo que contempla (i) añadir complejidad para poder probar el potencial del sistema, (ii) ser capaces de concluir cual es la manera más eficaz para llevarlo a cabo y (iii) eventualmente optimizar cada uno de los elementos implicados.





**Esquema 1. Diagrama de flujo representando las distintas fases del estudio** de funcionalidad de los objetivos del Modelo#1: sistema de intercambio de casete génica mediado por recombinasas (RMCE) y lentivirus deficientes en integración (IDLIV).

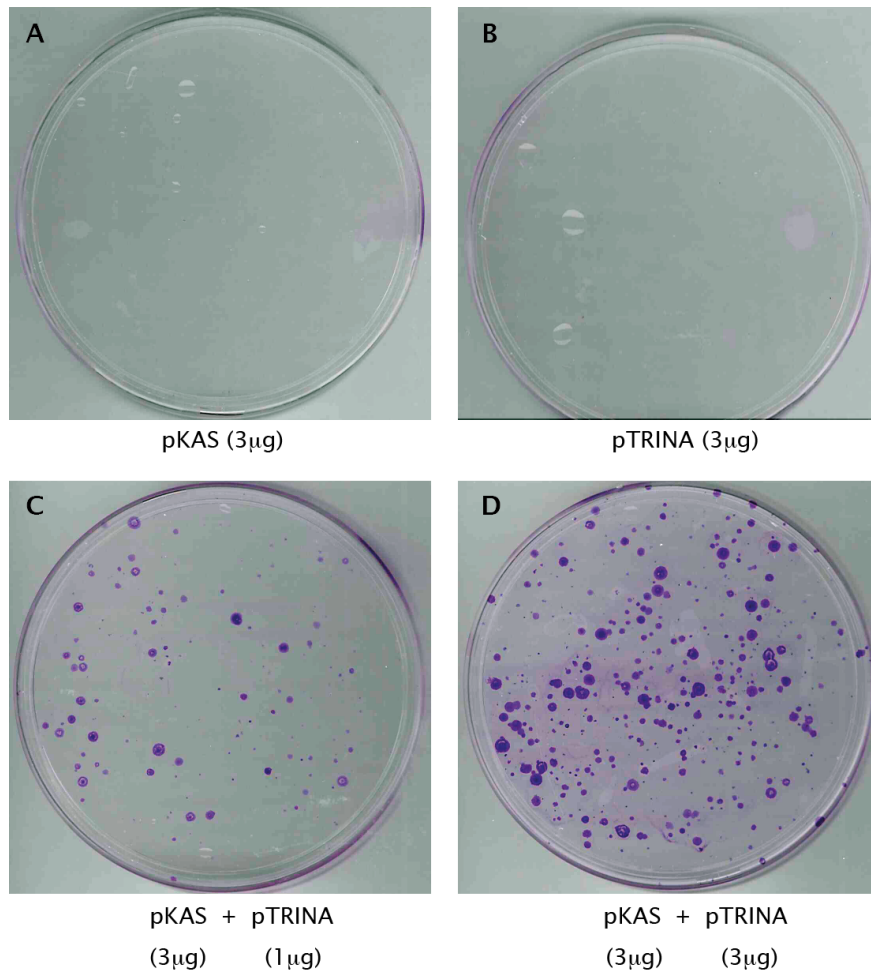
### 3. Prueba de función 1: RMCE en copias episomales

En una primera aproximación se testó la posibilidad de que dos sustratos episomales fueran capaces de recombinar, gracias a la acción de la recombinasa cre codificada en el plásmido pTRINA (figura 6). Se aprovechó el hecho de que los plásmidos lanzadera (para producir lentivirus) son funcionales en la célula en su formato plasmídico. Se realizaron transfecciones transitorias con los plásmidos pKAS o pTRINA, o co-transfecciones. La aparición de células resistentes a higromicina (higroR) y capaces de expresar la proteína roja fluorescente permitirá analizar la efectividad del sistema mediante ambos marcadores.

#### 3.1. Análisis de eventos positivos de recombinación por resistencia

Con el fin de rescatar clones higroR generados tras la reacción RMCE para su posterior análisis y caracterización, se dejaron transcurrir 72h tras la transfección; entonces los cultivos se diluyeron 1:8 y se dispusieron en placas de 10cm<sup>2</sup> con medio de selección de recombinación (DMEM-H, 100 µg/mL de higromicina) para aislar clones resistentes. Tras 10 días en cultivo, se fijaron y tiñeron para su cuantificación y análisis.

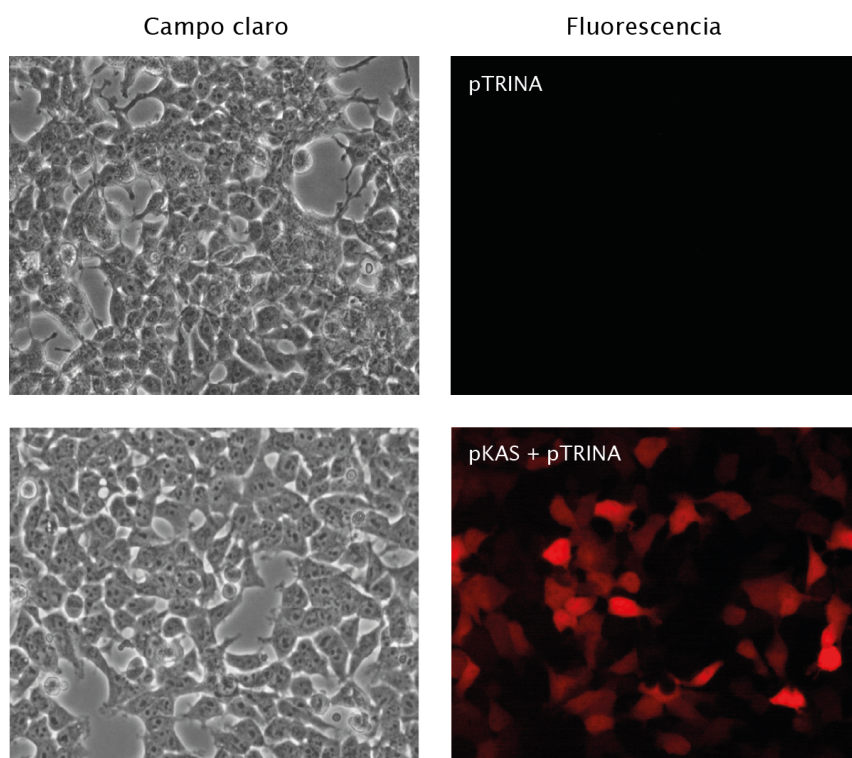
En los controles transfectados con los plásmidos pKAS o pTRINA, no se obtuvieron clones resistentes a higromicina (figura 7a y 7b, respectivamente). Sólo cuando se co-transfectaron pKAS y pTRINA, se obtuvieron clones de células higroR en una frecuencia que oscila entre  $1 \times 10^{-3}$  y  $3 \times 10^{-2}$  en relación a la cantidad de plásmido donador baja (1 µg) o alta (3 µg) que se emplee (figura 7c y 7d, respectivamente).



**Figura 7. Obtención de colonias higroR por RMCE entre copias episomales de DNA transfectadas.** Colonias resistentes a higromicina obtenidas la transferencia por transfección de los plásmidos indicados (pKAS / pTRINA). Los cultivos de HEK293A se tiñeron tras diez días de selección. Se indica por *low* y *high* condiciones de transfección a dosis relativa baja y alta respectivamente de pTRINA. Se muestra un ejemplo representativo de tres ensayos.

### 3.2. Análisis de eventos positivos por fluorescencia

Siguiendo el mismo protocolo que en el apartado anterior, transcurridas 72 horas tras la transfección se analizaron las células para comprobar si eran positivas para el marcador ChFP. El análisis mediante microscopía de epifluorescencia permitió valorar cualitativamente, que cuando se co-transfectaban pKAS y pTRINA se generaba un elevado número de células ChFP+ (figura 8). En cambio, y en consonancia con los datos obtenidos de colonias higroR, en los controles transfectados sólo con pKAS o pTRINA, no se detectaban células ChFP+.

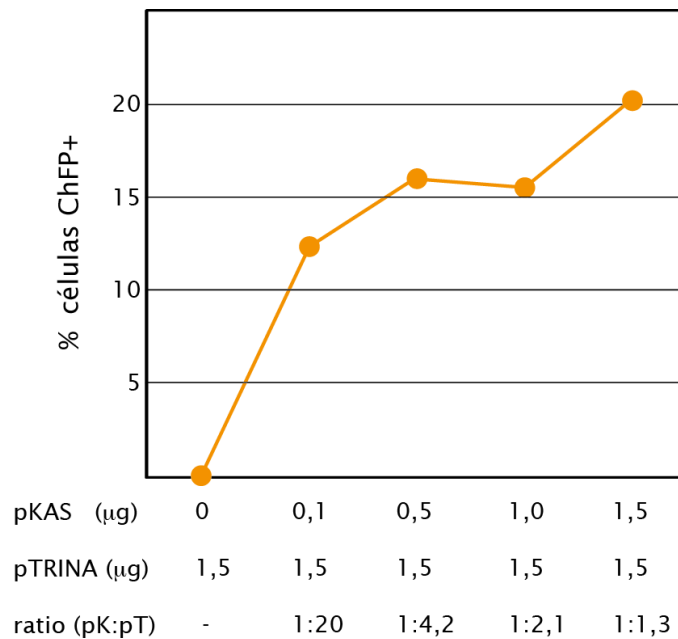


**Figura 8. Obtención de células ChFP+ por RMCE entre copias episomales de DNA transfectadas.** Presencia de células ChFP+ (rojas) como resultado de la reacción de RMCE en células transfectadas con pKAS (arriba) y pKAS /pTRINA (abajo). Se muestran microfotografías de la misma región tomadas por campo claro (izquierda) y por epifluorescencia (derecha).

Adicionalmente para ser capaces de cuantificar este fenómeno, se analizó mediante citometría de flujo la aparición de células ChFP+. Dados los resultados obtenidos con cantidades variables de pTRINA (comparación de figuras 7c y 7d), se optó por determinar si la relación del vector aceptor (pKAS) podía afectar a la frecuencia final de células ChFP+, es decir a la frecuencia de RMCE fidedigna; para ello se modificaron las cantidades relativas de plásmido aceptor, manteniendo fija la cantidad del donador (pTRINA).

En la figura 9 se muestran los porcentajes de células positivas sobre el total de células inicialmente transfectadas en relación a la cantidad absoluta de ambos plásmidos así como su ratio molar. Dichos valores oscilan entre el 12 y el 20%, aumentando al igualarse los ratios molares de los dos plásmidos. De nuevo los valores de los cultivos control no mostraron señal positiva alguna, como se observa en el primer punto de la gráfica, en el que sólo se utilizó pTRINA (dado que es el portador del gen que codifica para ChFP).

En conjunto estos resultados indican que cuando en una misma célula se da la presencia de ambos plásmidos, la recombinasa cre es capaz de catalizar la reacción y por tanto llevar a cabo el intercambio de casetes. Esto activa los elementos génicos rescatables mediante *promoter trapping*, expresando tanto los marcadores de resistencia como fluorescentes permitiendo la detección/selección de aquellas células en las que ha tenido lugar la RMCE.

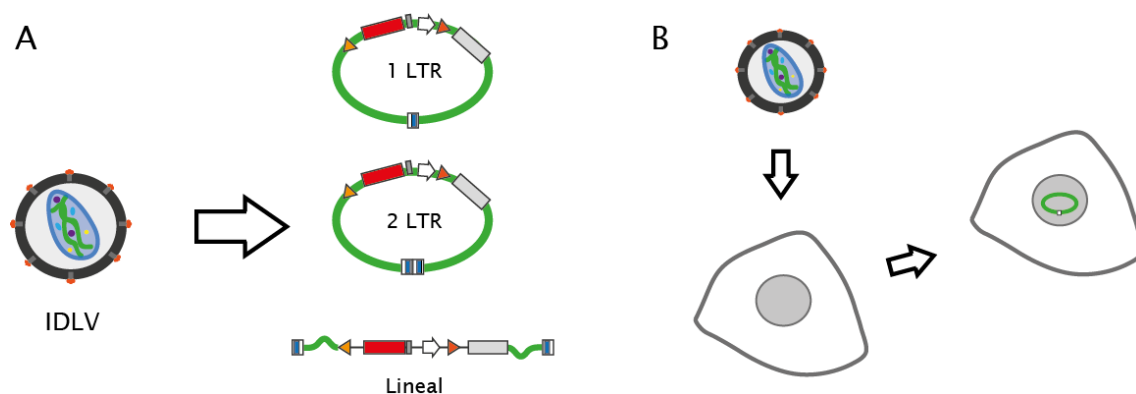


**Figura 9. Eficiencia de la reacción de RMCE entre plásmidos.** Número de células HEK293A que expresan ChFP en función de la cantidad de plásmido pKAS co-transfectado con una cantidad fija de pTRINA. Se indica la cantidad absoluta transfectada por placa de 10cm<sup>2</sup> de cada uno de los DNAs y la relación molar equivalente.

Por tanto, se puede concluir que la combinación de elementos diseñada es funcional y con una elevada frecuencia.

#### 4. Prueba de función 2.: Sistema de copias integradas

Se ha demostrado que dos sustratos episomales son capaces de recombinar entre sí gracias a la acción de la recombinasa cre codificada en uno de ellos. Por tanto, se propuso llevar a cabo dicha reacción entre un sustrato cromosómico (previamente integrado) y un cDNA lentiviral producido por la acción de la transcriptasa reversa, utilizando para ello IDLVs.



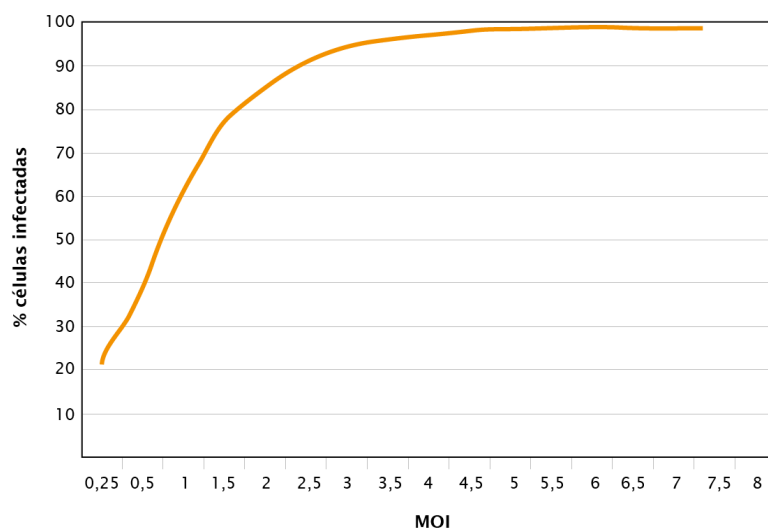
**Esquema 2. Representación esquemática** de (A) los diferentes productos de la reacción de retrotranscripción y (B) las consecuencias para la célula transducida con formas de cDNA episomales no integradas producto de vector deficiente en integración (IDLV).

#### 4.1. Sistema de entrega lentiviral. Generación y características

##### 4.1.1. Caracterización del vector lentiviral integrasa deficiente (IDLV)

Como se comentó en el apartado “Elementos del sistema”, la estrategia que se siguió está basada en el uso de vectores lentivirales integrasa deficientes. Con tal fin y conforme a lo descrito ampliamente en la literatura [136] se generó el plásmido pMDLg/pD64V que expresa los genes *gag-pol* de HIV en la producción de partículas recombinantes lentivirales con la mutación, en la integrasa lentiviral, que sustituye aspártico por valina en el aminoácido 64 (D64V).

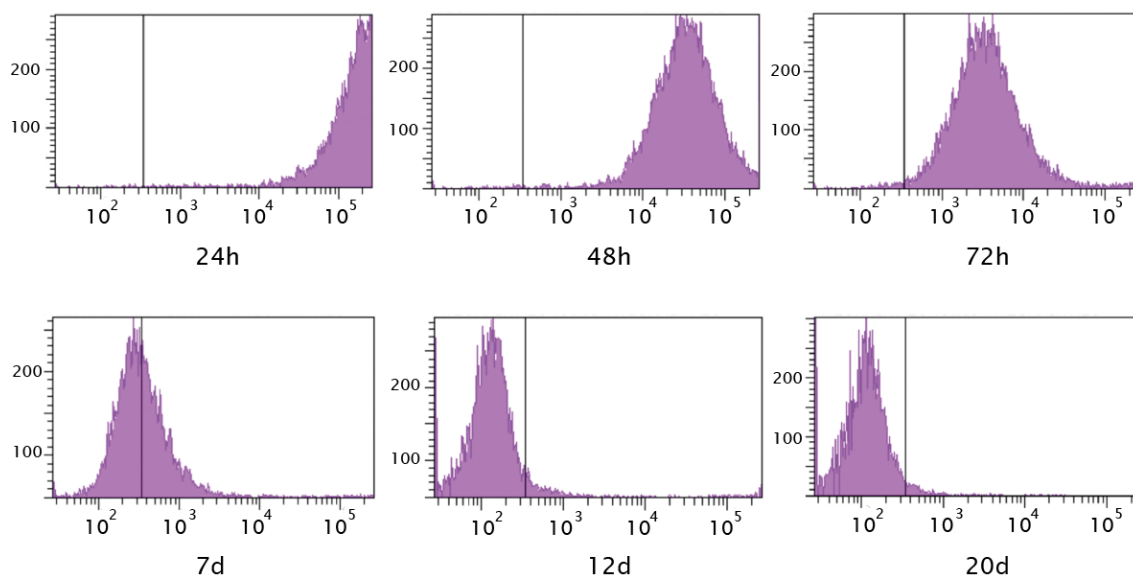
Para evaluar el fenotipo mutante (int-) se generaron lentivirus con esta forma D64V (IDLV) que expresan la proteína verde fluorescente (eGFP) bajo el control del promotor CMV (IDLV-CMV.eGFP) y se transdujeron células HEK293A a una MOI de 5 TU/cel, lo que permite la transducción de la mayor parte de las células (aproximadamente 95% de la población celular en este sistema [193]) (gráfica 1).



**Gráfica 1.** Representación de la variación del número de célula transducidas (infectadas) en un cultivo permisivo en función de la multiplicidad de infección empleada (MOI).

Asimismo, se generaron vectores control LV-CMV.eGFP y se estudió comparativamente la cinética de mantenimiento de células eGFP+ en los cultivos transducidos con éste y con el IDLV-CMV.eGFP. Como se aprecia en la figura 10, a tiempos tempranos tras la transducción (24, 48, 72 hpt), los virus IDLV (int-) se comportan de la misma manera que los virus salvajes, int+. Sin embargo tras la expansión del cultivo y las sucesivas divisiones celulares (7, 12, y 20 dpt), la proporción de células eGFP+ disminuye progresivamente en

los cultivos transducidos con IDLVs, y a 20 dpt alcanza valores idénticos a los de los cultivos control no transducidos.



**Figura 10. Análisis de FACS** de la evolución en el tiempo de la expresión de eGFP tras la transducción de células HEK293A con virus IDLV-CMV.eGFP . Se indican las horas transcurridas tras la transducción (h) o los días (d).

Por tanto, se generaron lentivirus defectivos en integración que funcionan según lo anticipado, ya que la pérdida de su expresión es coherente con su carácter no integrativo.

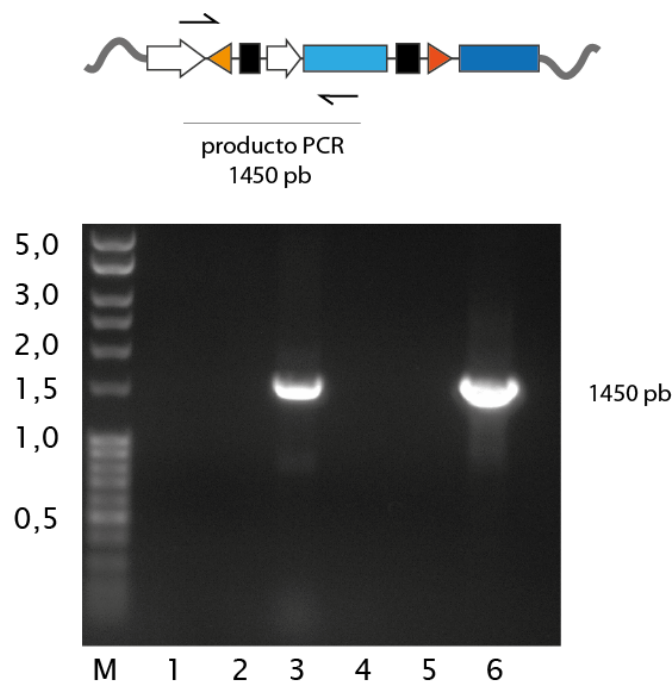
## 4.2. Generación de la línea estable HEK293 KAS

### 4.2.1. GLP (genetic landing pad): Sustrato cromosómico

Para generar la línea celular que porte el plásmido aceptor en un entorno cromosómico o sitio de posicionamiento genético (GLP, del inglés *genetic landing pad*) se realizaron transfecciones sobre células HEK293A con el plásmido pKAS, que contiene el gen *neo* dentro de la casete de intercambio (figura 6) o con un plásmido control (pCMV.eGFP).

Los cultivos transfectados se mantuvieron en medio de selección (DMEM-G, 450 ug/mL de G418) durante 14 días y se obtuvieron más de 80 colonias G418<sup>R</sup> aisladas que se reunieron en un cultivo único (*pool*), el cual fue expandido y seleccionado durante 7 días adicionales para eliminar posibles trazas de células negativas.

Se procedió a analizar genotípicamente, mediante PCR, el pool celular para confirmar que dichas células tenían integrada la casete KAS. Se utilizaron parejas de *primers*, uno *forward* específico de la región externa al casete (PGK fw, figura 11), pero contenida en el cDNA del vector lentiviral y otro *reverse* que hibridaba en el gen que confiere la resistencia a G418 (*neo rv*, figura 11). El resultado confirmó que el fenotipo observado correlaciona con la presencia de copias integradas de KAS.



**Figura 11. Análisis genotípico de la reacción de RMCE en copias integradas al azar (I).** El esquema indica la posición de los *primers* empleados y el tamaño del producto de PCR de la casete KAS en el DNA genómico de los cultivos. Se muestra el resultado obtenido por PCR convencional sobre el DNA genómico extraído de los cultivos de HEK293A parentales (carril 1) o transducidas con el LV control LV-CMV.eGFP (carril 2). Los carriles 3 al 5 muestran el resultado sobre el *pool* de células neo<sup>R</sup> parentales (293AKAS, carril 3) o transducidas con IDLV.TRINA a MOI 2 y 10 (carril 4 y 5 respectivamente). El carril 6 muestra el producto obtenido con el plásmido control pKAS.

### 4.3. RMCE mediado por IDLV

Se han generado y caracterizado los elementos necesarios para desarrollar el modelo, es decir, tanto IDLVs como células diana con el casete aceptor integrado. Por lo que se realizaron los ensayos de transferencia mediante transducciones sobre las células caracterizadas.

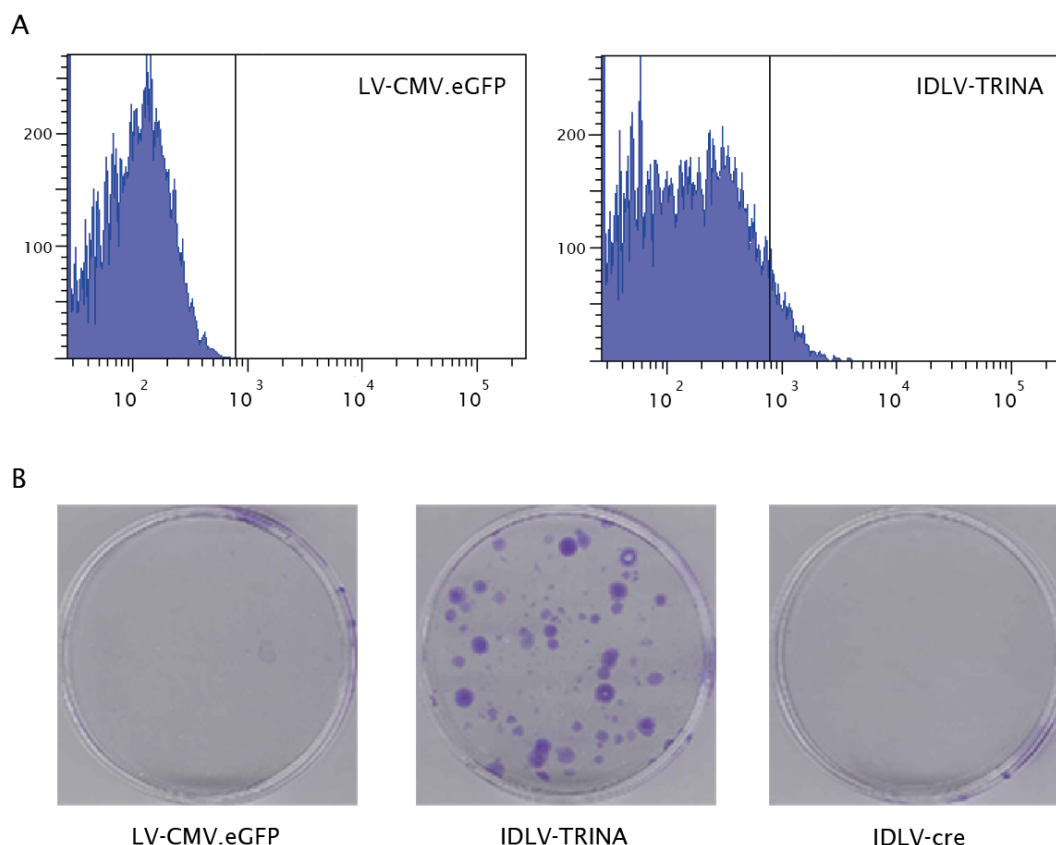
El modelo permite la expresión de la recombinasa cre desde el módulo de transferencia TRINA (ver “Elementos del diseño. Expresión limitada”). Dado que la MOI condiciona el promedio de partículas lentivirales por célula y por tanto el número de cDNAs virales por célula (1 molécula de cDNA por partícula transductora); ha de jugar un papel importante tanto en la eficiencia de la reacción como en la extensión de la genotoxicidad de la enzima.

En un primer ensayo se evaluó la funcionalidad del sistema a una MOI baja (2 TU/cel), que pudiera asegurar una transducción celular razonable (70-80%, gráfica 1) y así promover recombinación en el mayor número de células posible, pero evitando posibles problemas de toxicidad asociados a una alta expresión de la recombinasa. El experimento se diseñó para permitir dilucidar si las colonias resistentes a higromicina (higro<sup>R</sup>) son debidas a eventos de recombinación fidedignos, o a eventos de delección erróneos entre los sitios *loxP*



heteroespecíficos contenidos en KAS. En este segundo caso, las colonias no podrían ser positivas para la proteína roja fluorescente (ChFP+).

Se transdujeron 500.000 células HEK293A a MOI 2 con un lentivirus control (LV-CMV.eGFP), con el lentivirus objeto de estudio (IDLV-TRINA) y con un lentivirus que carece de casete para el intercambio, pero que codifica para la expresión de la recombinasa cre (IDLV-cre) (figura 12). A 5dpt los cultivos se expandieron 1:8 en placas de 10cm<sup>2</sup>, manteniéndose durante diez días en medio DMEM-H diluidos para poder obtener clones aislados y fácilmente cuantificables.



**Figura 12. Análisis fenotípico de la reacción de RMCE en copias integradas al azar (I).** (A) Proporción de células ChFP+ presentes en los cultivos de 293KAS transducidos con vector control (LV-CMV.eGFP) o con IDLV-TRINA tras los 5dpt analizado por FACS. (B) Estudio de la frecuencia de aparición de colonias higro<sup>R</sup> tras la transducción de células 293KAS con IDLV control (IDLV-CMV.eGFP), con un IDLV que sólo expresa cre (IDLV-cre) o con el vector TRINA (IDLV-TRINA). Los cultivos fueron transducidos a MOI de 2 TU/cel. En B, se mantuvieron durante diez días en selección antes de revelar el número de colonias aparecidas.

Únicamente cuando se aportó tanto la enzima como el casete de intercambio TRINA se obtenían tanto colonias ChFP+ (IDLV-TRINA, figura 12a) como colonias higro<sup>R</sup> (IDLV-TRINA, figura 12b) mientras que cuando el virus utilizado era un control (LV-CMV.eGFP, figura 12a y 12b, paneles izquierda) o poseía la enzima pero no el casete de intercambio (IDLV-cre, figura 12b), no se obtenían ni colonias resistentes, ni células ChFP+.



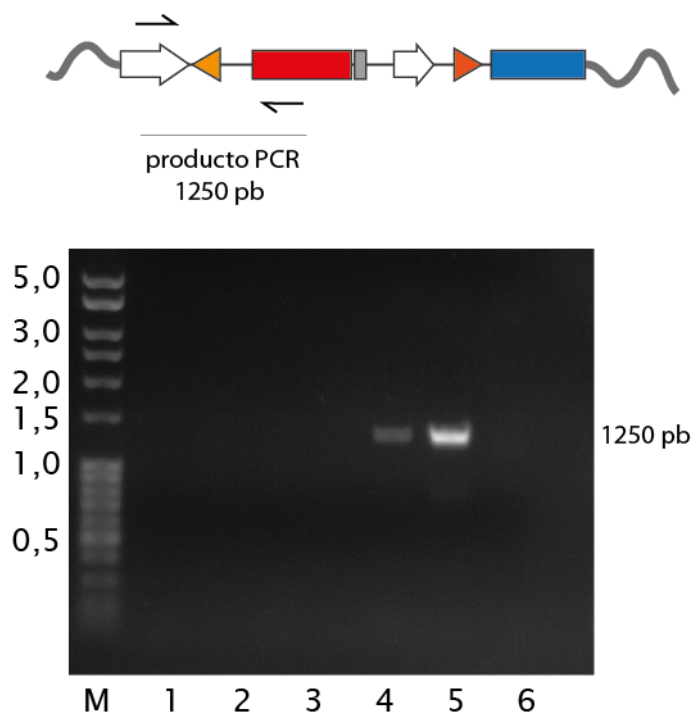
Por tanto, se puede concluir que la estrategia diseñada de RMCE es funcional y viable cuando se disponen de copias de la casete aceptora integradas al azar en el genoma celular. Además, la entrega de la casete donadora a través de IDLVs resulta eficaz para los procesos de RMCE.

#### 4.3.1. Caracterización genotípica de células positivas tras la reacción de RMCE

Se caracterizaron genotípicamente las células obtenidas para determinar si la recombinación había tenido lugar en la casete aceptora o si pudiera tratarse de eventos de integración no específicos (negativos) que rescataran el fenotipo observado en el pool KASTRINA. Para ello, se diseñó un ensayo de PCR que permitía dilucidar los eventos de recombinación fidedignos, mediante el uso de una pareja de *primers* que hibrida tanto en el DNA codificante para la ChFP (ChFP rv, en el interior de la casete intercambiada) y en el promotor que dirigirá su expresión tras la recombinación (PGK fw, contenido en el casete aceptor integrado inicialmente). Únicamente en los casos en los que la reacción haya ocurrido de manera precisa se obtendrá una banda de 1250pb.

Como controles de la reacción de PCR se utilizaron DNAs de las células transducidas con lentivirus control (LV-CMV.eGFP), de las células HEK293A.KAS sin transducir y de células parentales HEK293A sin ningún tipo de modificación. Sólo en el caso de la línea celular transducida con el IDLV-TRINA se obtiene la banda esperada de 1250 pb (figura 13, carriles 5 y 6), que nos demuestra la recombinación fidedigna. Mientras que ni en la línea parental (figura 13, carriles 2 y 3), ni en las células transducidas con el lentivirus control (figura 13, carril 3), se obtiene dicha banda.

Esto indica que la reacción ha procedido en la forma anticipada y que tras un período de 20 días de selección con higromicina, sólo se rescatan células en las que se ha llevado a cabo el proceso de recombinación de forma fidedigna (figura 11 y 13).



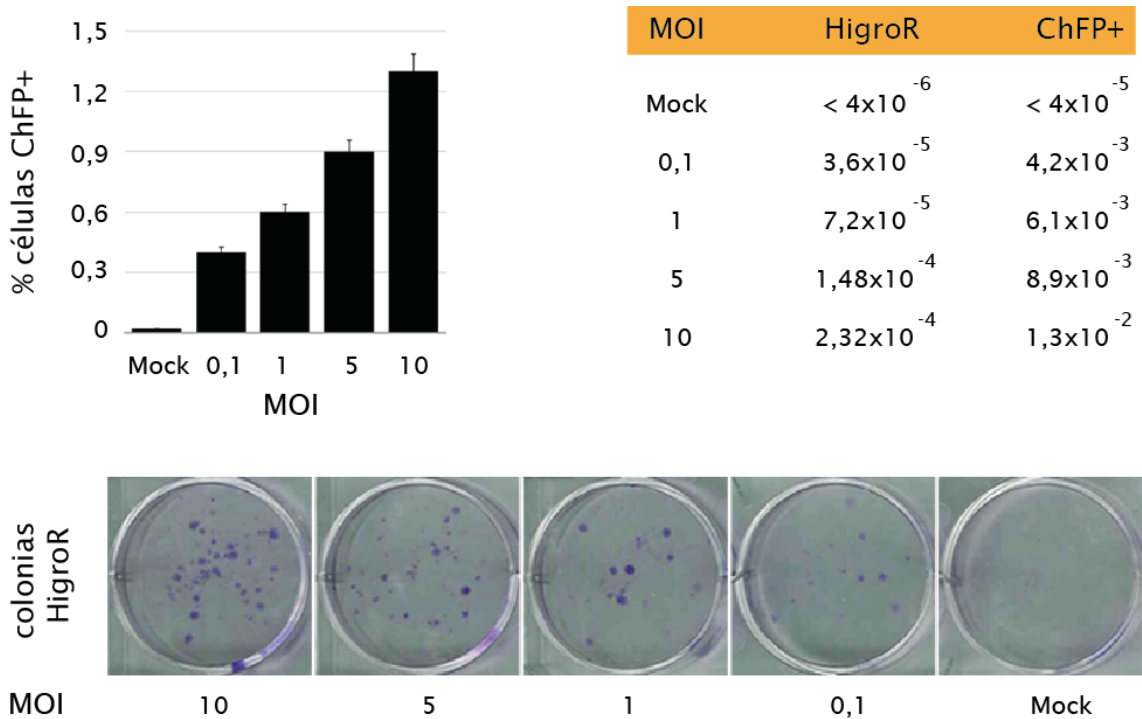
**Figura 13. Análisis genotípico de la reacción de RMCE en copias integradas al azar (II).** El esquema indica la posición de los *primers* empleados y el tamaño del producto de PCR de la casete TRINA en el DNA genómico de los cultivos. Se muestra el resultado obtenido por PCR convencional sobre el DNA genómico extraído de los cultivos: HEK293A parentales (carril 1); transducidas con el virus control LV-CMV.eGFP (carril 2); el *pool* de células neo<sup>R</sup> parentales 293AKAS (carril 3); transducidas con IDLV.TRINA a MOI 2 y 10 (carril 4 y 5 respectivamente); el carril 6 con el plásmido control pKAS.

#### 4.3.2. Optimización del proceso de RMCE mediado por IDLV

Puesto que el proceso de RMCE es funcional sobre copias integradas utilizando IDLVs, y dado que se ha mencionado el papel que la MOI puede jugar en este proceso; se planteó la posibilidad de optimizarlo para encontrar las condiciones (de MOI) que maximizaran la recombinación haciéndola más eficiente, es decir, permitiendo obtener un mayor número de colonias positivas. Para ello se utilizaron distintas multiplicidades de infección que van desde 0,1 hasta 10 TU/cel. Los datos se cuantificarán como la frecuencia relativa de *targeting* efectivo (ETF) medida como el número de células fluorescentes (ChFP+) o colonias higroR que se obtienen a partir del número de células inicialmente transducidas.

Tomando los datos obtenidos (figura 14), se realizó la tabla 3 (embebida en la figura 14), que muestra como los valores oscilan entre  $2,3 \times 10^{-4}$  (MOI 10) y  $3,6 \times 10^{-5}$  (MOI 0,1) en el caso de las colonias resistentes para higromicina y entre  $1,3 \times 10^{-2}$  (MOI 10) y  $4,2 \times 10^{-3}$  (MOI 0,1) en el caso de las células que se cuantificaron por FACS. Si los valores obtenidos de ETF a baja MOI (1 y 0,1 TU/cel) se corrigen por la probabilidad de transducción de acuerdo a la distribución de Poisson [194], ( $p(\text{transducción}) < 0,67$  y  $< 0,099$  respectivamente) las frecuencias de recombinación son entonces valores de la misma magnitud independiente de la MOI, tanto para ETF medida como higroR:  $1,08 \times 10^{-4}$  (MOI

10) y  $3,6 \times 10^{-4}$  (MOI 0,1); como ETF medido como ChFP+:  $\approx 1 \times 10^{-2}$  (MOI 1) y  $4 \times 10^{-2}$  (MOI 0,1). Ello sugiere que la eficiencia de la reacción de recombinación está limitada por la probabilidad de transducción como reflejo de la MOI.



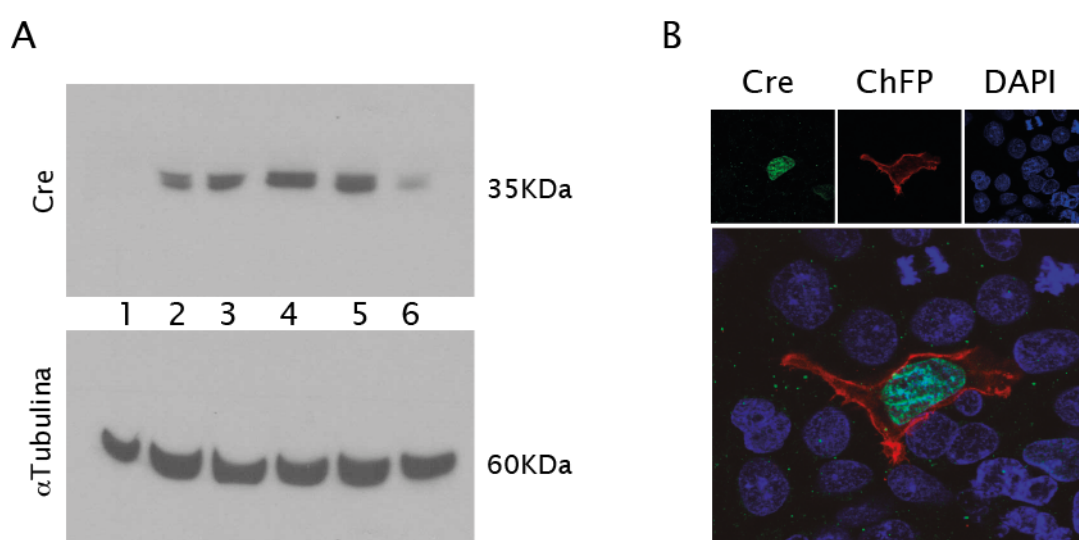
**Figura 14. Efecto de la MOI en la frecuencia de la RMCE mediada por IDLV-TRINA.** En el panel superior izquierda, se indica el porcentaje de células ChFP+ de los cultivos de 293KAS transducidos a la MOI indicada de IDLV-TRINA y analizados por FACS a las 72 hpt. En el panel inferior se presentan los datos obtenidos a los 10dpt analizados como numero de colonias higroR obtenidas. Se indica la frecuencia como ETF (*efficient transduction frequency*) calculada como la proporción de clones obtenidos en cada condición respecto al número de células sembradas antes de la transducción. No se ha corregido con la probabilidad teórica de transducción, según distribución de Poisson (ver texto). La tabla superior derecha, resume los datos recogidos en los gráficos.

Cuando se transduce el virus control (Mock o LV-CMV.eGFP) no se observó ninguna colonia, ni célula ChFP+ y los valores de ETF son  $<10^{-6}$  y  $<10^{-5}$  (higroR o ChFP+ respectivamente) lo que indica que los eventos de recombinación no específica usando este sistema son altamente improbables y al menos 2 órdenes de magnitud inferiores a las frecuencias obtenidas mediante RMCE.

Estos resultados sugieren que el empleo de un sistema viral de transferencia para la expresión de la recombinasa es de gran utilidad; dado que dependiendo del tipo celular se podrían adaptar las condiciones de transducción y así optimizar en cada caso particular eficiencia o eficacia de la reacción de RMCE.

### 4.3.3. Verificación de la expresión limitada de cre

Los resultados anteriores sugieren que los niveles auto-limitantes de cre (recordar que cre está codificada dentro del sistema donador IDLV-TRINA, figura 6), son el cuello de botella de la reacción. Se estudió mediante *Western blot* la cinética de expresión de la recombinasa cre (desde el IDLV-TRINA) a distintos tiempos post-transducción con una MOI constante de 10 TU/cel. La figura 15 muestra como los niveles de cre aumentan entre las 12hpt (carril 2) y las 72hpt (carril 5) y comienzan a disminuir hasta niveles casi indetectables a los 7dpt (línea 6). Resultados similares se obtuvieron mediante inmunofluorescencia de células con un anticuerpo específico (ver figura 15b, se muestra resultado a las 24hpt). Es destacable que la ausencia de expresión de cre en los cultivos no-transducidos apoya la idea de que la expresión de cre codificada en los IDLV es necesaria para la obtención de células con fenotipo higoR y ChFP+.



**Figura 15. Expresión de cre por IDLV-TRINA.** (A) Se muestra la cinética de expresión de cre obtenido por *Western blot* de células 293KAS transducidas con IDLV-TRINA. Los cultivos se analizaron a las 12, 24, 48, 72 y una semana post-transducción. En el carril 1 se muestra los resultados de células control no transducidas. (B) Imagen de inmunofluorescencia del cultivo transducido con IDLV-TRINA a las 24hpt (centro, ChFP+) revelado con anticuerpo específico de cre (izquierda, verde) y los núcleos marcados con DAPI (derecha, azul). Se muestra el *merge* en la imagen principal.

## 5. Prueba de funcionalidad 3: RMCE sobre copia única integrada en sitio específica en HEK293A

El siguiente ensayo de funcionalidad se planteó sobre secuencias aceptoras presentes como copia única en un locus específico del genoma. Esta estrategia persigue la actuación sobre un locus definido que permitiera el mínimo efecto sobre la función de los genes adyacentes y admitiera la expresión de la casete sin restricciones asociadas a modificaciones de la cromatina.

## 5.1. Generación del GLP en un locus determinado

### 5.1.1. Transferencia por recombinación homóloga. Locus humano ROSA26

La recombinación homóloga (acrónimo del inglés, HR) es el procedimiento tradicionalmente empleado para insertar en un locus concreto una secuencia determinada. Por razones poco conocidas y descritas someramente en la introducción, las células humanas muestran una baja predisposición a este fenómeno. En realidad el trabajo de Keller y colaboradores[195] aventuraba un locus con especial predisposición a ser modificado por HR al igual que ocurre en el ortólogo murino ROSA26. Debido a ello se eligió este locus como diana de inserción de KAS entre los nucleótidos 9.415.082 y 9.414.043 en el cromosoma 3 (GenBank:CR624523) que son los empleados en la HR en el ratón.

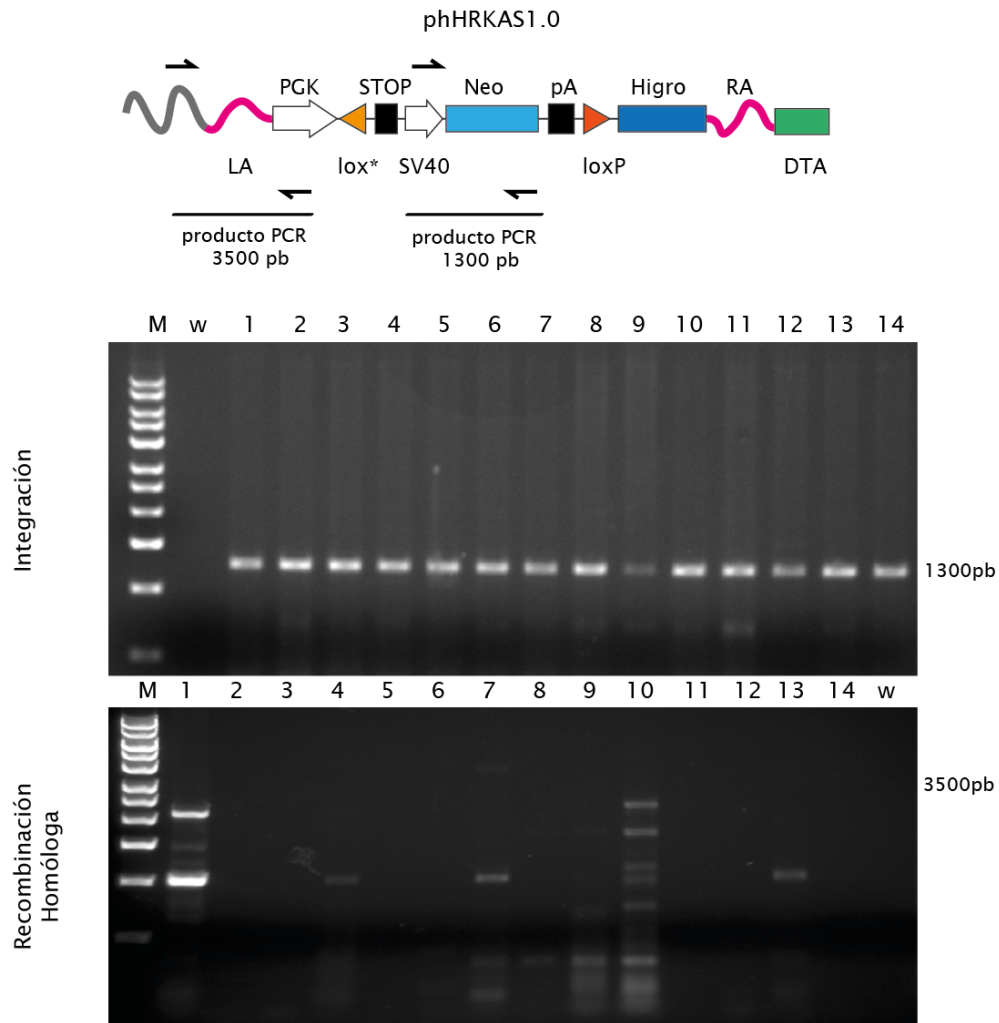
Para generar el vector de recombinación se partió del plásmido pKAS1.0 empleado anteriormente, sobre el que se clonaron los brazos de homología necesarios para llevar a cabo la reacción de recombinación homóloga (figura 16). Dichos brazos se componen de un fragmento genómico de 3Kb en el extremo 5' (LA) y otro de 2Kb en el extremo 3' (RA). Finalmente en este extremo se clonó como marcador de selección positivo el gen que codifica para la toxina diftérica (figura 16) generándose el vector resultante que se denominó phRKAS1.0.

Células HEK293A se transfectaron, con el plásmido linearizado con EcoRV, tanto por métodos químicos clásicos, como mediante nucleofección. Transcurridas 48 horas, los cultivos se cambiaron a medio de selección (DMEM-G) durante dos semanas para poder aislar colonias en las que el casete se hubiera integrado de manera estable. Debido a la frecuencia descrita para este locus (1:24) [195] se propuso el análisis de 30 clones aislados. Mediante el empleo de *primers* específicos externos a los brazos de homología junto a otros internos permitió determinar que ninguno de los clones obtenidos había integrado el plásmido phRKAS1.0 por HR si bien las células contenían dicha construcción insertada (figura 16<sup>1</sup>).

Dada la incapacidad de poder obtener clones con la casete aceptora insertada específicamente en el locus ROSA26, se decidió optar por un abordaje distinto que pudiera promover la obtención de clones con la copia en el locus de elección.

---

<sup>1</sup> Se puede apreciar que sólo se muestra el resultado de 14 de los clones, si bien son 30 los analizados; pero se decidió de esta manera para no aportar mayor complejidad a la figura.



**Figura 16.** Estructura de los elementos contenidos en el plásmido de recombinación homóloga específico para ROSA26 humano. Se indican los elementos básicos contenidos en el plásmido phHRKAS1.0: promotor, rectángulo apuntado; ORFs, rectángulos; señales lox, triángulos; y secuencia de terminación repetida como caja. Los brazos de homología se representan como líneas curvas de color rosa. En la parte inferior se representa el análisis por PCR sobre DNA genómico de 14 de los 30 clones aislados y crecidos en selección con G418 (panel inferior). Se indica la posición de los *primers* empleados sobre el esquema de la parte superior y el tamaño esperado. Como control se muestra el gel de agarosa con los productos de PCR obtenidos con *primers* que hibridan dentro de la secuencia del ORF de neo (panel superior). (w) indica reacción control sin DNA.

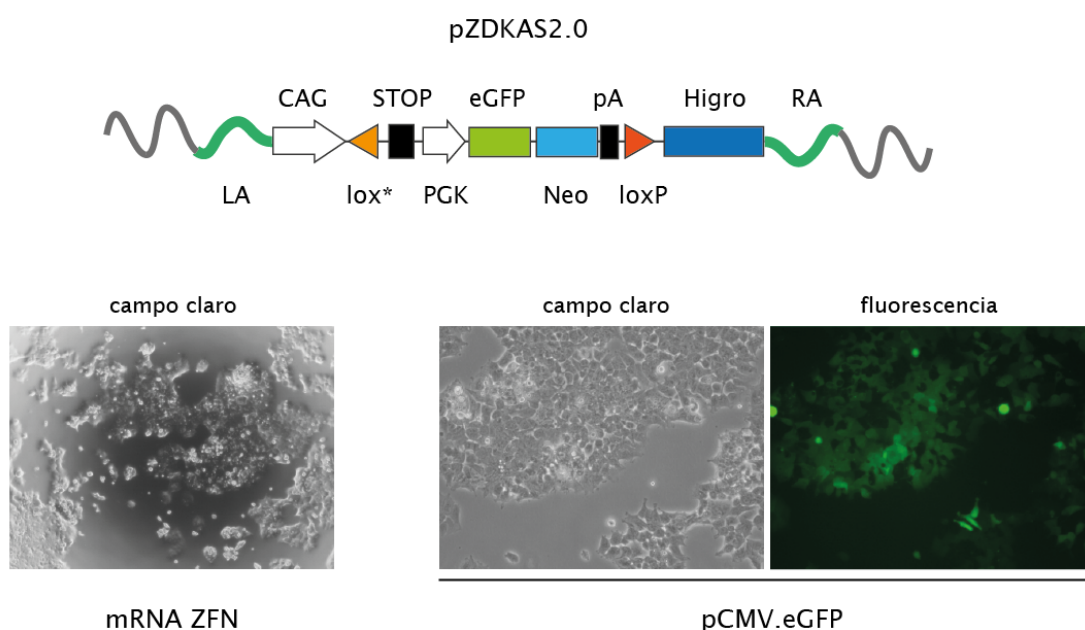
### 5.1.2. Transferencia por recombinación homóloga con ZFN. Locus AAVS1

Dada la dificultad observada en la obtención de células humanas por HR, se decidió cambiar la estrategia por un aproximación más efectiva, la cual hiciera uso de la tecnología derivada de las nucleasas de dedos de zinc (ZFN). Como se comentó en la Introducción estas nucleasas promueven la HR tras la inducción de DSB (rotura de doble cadena) en sitio específica, elevando la tasa de eficiencia muy por encima del método convencional[75]. La acumulación de datos positivos en su empleo [63], [196] y la disponibilidad comercial durante el desarrollo de la presente tesis hizo que se optara por elegir el locus *AAVS1* (19q13.4-qter). Este locus está considerado como sitio inocuo, en el cual se producen de

manera natural las inserciones de virus adeno-asociados[56]. Además, su haploidía no produce efectos adversos en sistema tan delicados y complejos como las células embrionarias humanas [197].

La casete donadora pKAS2.0 se insertó en el plásmido de recombinación pZDonor. El vector resultante pZDKAS2.0 (figura 17) se utilizará para llevar a cabo el proceso de integración sitio específica, mediante la utilización de la ZFN correspondiente.

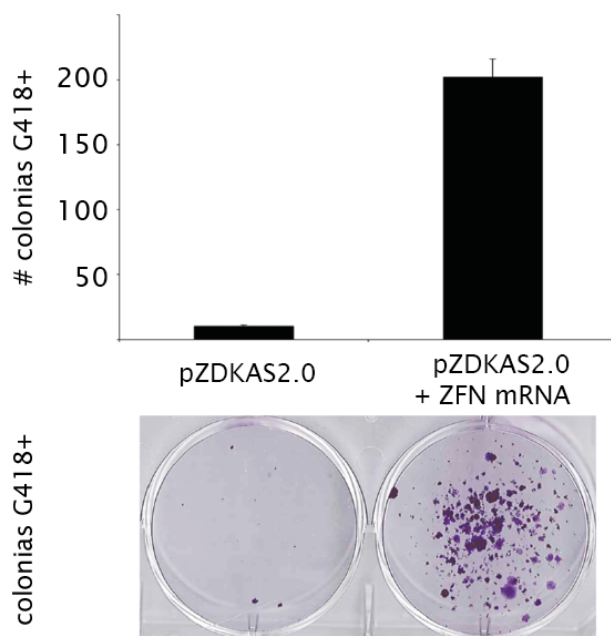
Mediante transfección transitoria, sobre células HEK293A, se introdujeron los distintos elementos, utilizando los controles oportunos: (i) un vector reportero pCMV.eGFP que servirá como control de la eficiencia de transfección, (ii) el plásmido pZDKAS2.0 (sin enzima) que ofrecerá la frecuencia de integraciones aleatorias (iii) el mRNA que codifica para la nucleasa, como valoración de la toxicidad para la célula y (iv) la co-transfección del plásmido pZDKAS2.0 y el mRNA de la nucleasa ZFN. A las 24hpt se analizó tanto la viabilidad celular (en particular en los cultivos que estaban transfectados con el mRNA de la ZFN) como el porcentaje de células eGFP+. Mientras que el uso del mRNA tiene un efecto citotóxico débil pero consistente, que se manifestó por la presencia de células en suspensión; la tasa de transfección fue alta (figura 17).



**Figura 17. Estructura de los elementos contenidos en el plásmido específico para recombinación homóloga en AAVS1 mediado por ZFN.** (A) Esquema de la organización de los componentes contenidos en el plásmido pZDKAS2.0 para inserción mediada por nucleasa de dedos de zinc (ZFN). Se ha mantenido la misma leyenda que en la figura anterior para indicar elementos de igual naturaleza funcional. Se ha omitido por claridad indicar P2A en la casete bi-cistrónica eGFP-neo. Los brazos de homología se representan por bloques verdes. (B) Microfotografías tomadas a las 24hpt de campo claro (izquierda y centro) o epifluorescencia (derecha) de los cultivos transducidos con sólo el mRNA de la ZFN específica de AAVS1 (izquierda) o co-transfectada con el plásmido de recombinación pZDKAS2.0.



Una vez valorados estos datos, que arrojan un balance favorable en la consecución del experimento, se añadió medio de selección (DMEM-G). Diez días después se tiñeron los cultivos y se cuantificó la eficiencia de generación de clones. La combinación de pZDKAS2.0 y la nucleasa permitió obtener 100 veces más clones neoR que en el control sin ZFN (pZDKAS2.0) (figura 18). Se aislaron y amplificaron más de 50 clones, de los cuales 12 se caracterizaron genotípica y fenotípicamente.



**Figura 18. Eficiencia de rescate de colonias G418R por ZFN.** Se muestra el número (arriba) y el aspecto de la placa teñida (abajo) de colonias G418R tras la transfección con los plásmidos indicados con o sin ZFN de AAVS1. Se representan los resultados de uno de tres ensayos realizados.

#### 5.1.2.1. Caracterización genotípica de clones 293A KAS2.0

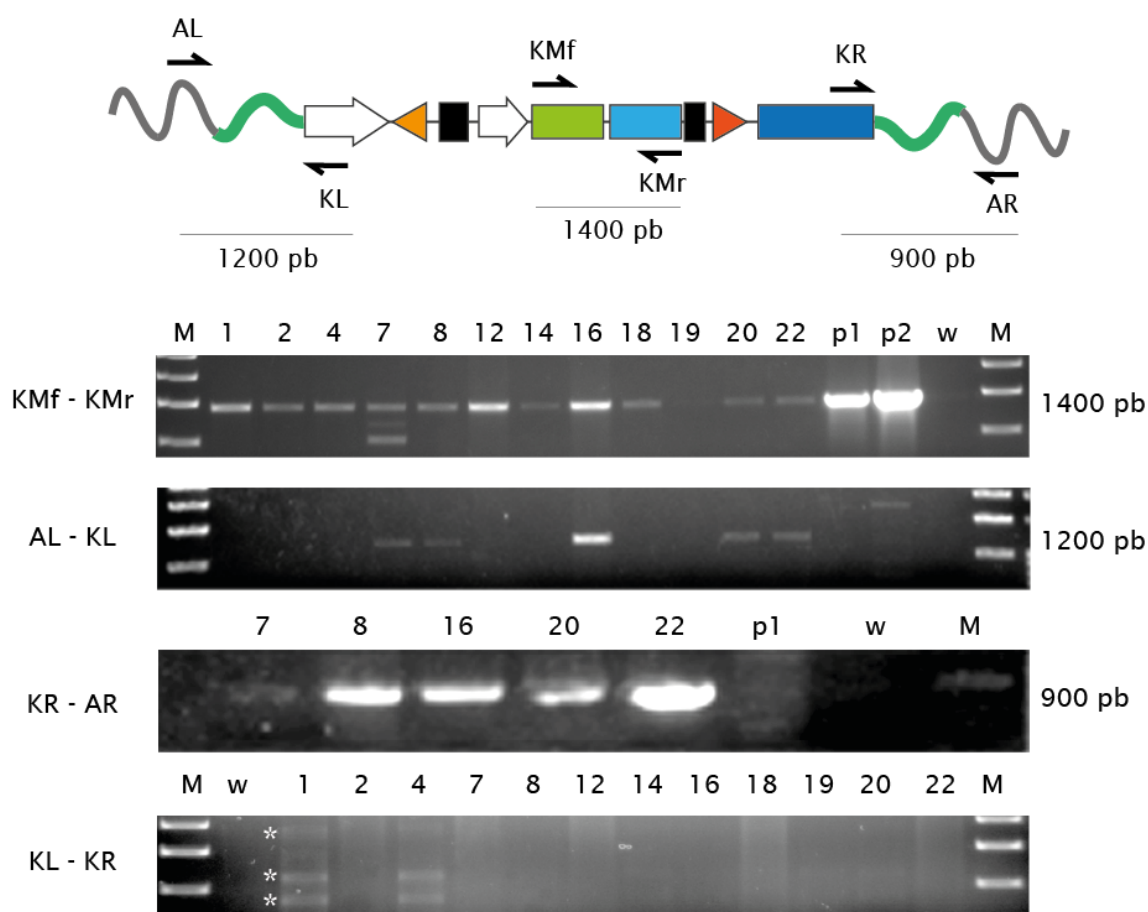
##### a) Análisis mediante PCR

Los clones aislados fueron analizados por PCR sobre DNA genómico obtenido de los diferentes cultivos; primero para comprobar la integración de la casete mediante una pareja de *primers* internos a la construcción KAS2.0 (KMf / KMr, específicos de integración) y posteriormente mediante una combinación de un primer externo (situado en el locus cromosómico) y uno interno (casete) se detectaron procesos de recombinación homóloga (AL / KL, en el caso del brazo de homología izquierdo y KR / AR en el derecho).

Como se detalla en la figura 19 la presencia de una banda de 1,4Kb (figura 19 foto superior), indicó que los doce clones habían integrado el casete en su genoma, de los cuales un 17% (5 de 12), los clones #7, #8, #16, #20 y #22 tienen el casete en el sitio específico deseado, como se puede apreciar a partir de las bandas de 1,2Kb y 0,9Kb (figura 19 centro arriba y 19 centro abajo, respectivamente).



Para poder detectar los clones que además de tener integraciones en sitio específico hubieran integrado una única copia se utilizó la pareja de *primers* de los extremos (KL / KR), para ver cuales de los clones poseían más de una copia integrada en el sitio *AAVS1* o lo que se define como concatémeros *head-to-tail* (cabeza-cola). Éstos productos, se generan por reacciones parciales de recombinación homóloga de sólo uno de los extremos y/o mediado por recombinación no homóloga entre los extremos o *Non Homologous End-Joining* (NHEJ). Se detectó que los clones #1 y #4, ambos con el casete integrado pero no en el locus *AAVS1*, habían sufrido posiblemente un proceso de concatemerización, ya que aparece tanto la banda de 2Kb (concatémeros *head-to-tail*) como las bandas de 1.1Kb y 0.9Kb (posibles reacciones parciales de NHEJ). Mientras que los clones #7, #8, #16, #20 y #22 sólo poseen una copia de la casete en el sitio deseado (figura 19 abajo).



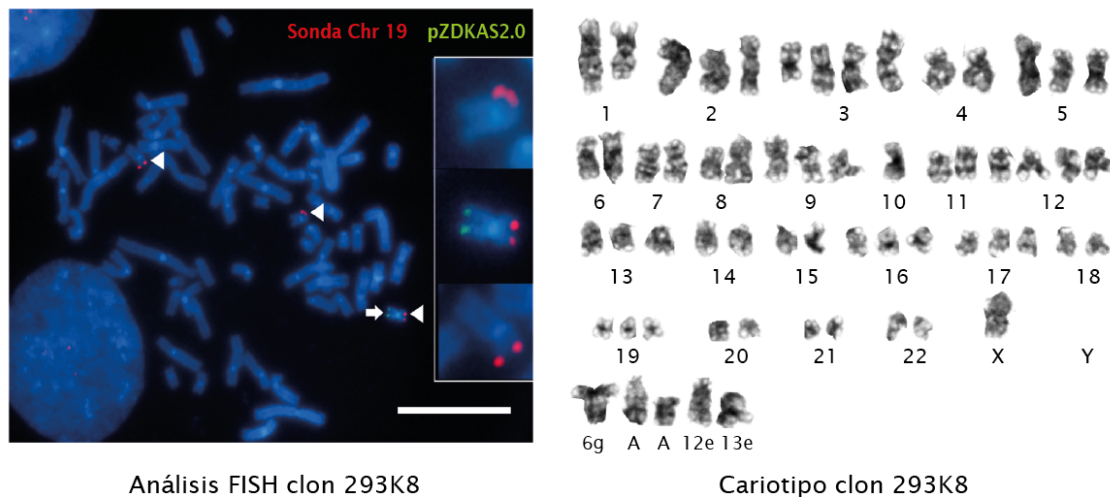
**Figura 19. Caracterización genotípica de clones 293A KAS2.0 (PCR).** Se muestra el esquema de la casete integrada en el locus *AAVS1* y la posición de los *primers* empleados en su caracterización. Los paneles inferiores muestran los resultados de PCR obtenidos sobre DNAg de 12 clones G418R aislados de los 24 estudiados. La selección de *primers* en cada ensayo (izquierda) permite detectar presencia (foto superior), integración por HR en el locus *AAVS1* (las dos centrales) o integraciones concatéméricas (foto inferior) (ver texto para detalles) Los asteriscos indican la presencia de formas generadas por concatémeros. M. Marcador de peso molecular.

## b) Análisis mediante FISH

Se quiso confirmar en los clones con integración en AAVS1 si ésta era única y en dicho locus. Para ello se seleccionaron dos clones #293K8 y #293K16 que fueron estudiados por hibridación de fluorescencia *in-situ* (FISH) mediante el empleo de (i) una sonda específica, que permitió ver en que cromosomas aparecen integradas copias de pZDKAS2.0 y (ii) una genérica, control del brazo corto del cromosoma 19, que permitiera detectarlo sin interferir con la señal del sitio AAVS1.

Consistente con el cariotipo hipotriploide descrito para esta línea celular (ATCC CRL-1573, y cariotipo en figura 20) en ambos clones estudiados #293K8 y #293K16 la sonda genérica (rojo en figura 20), marca tres copias del cromosoma. Mientras que la sonda específica para pZDKAS2.0 (verde en figura 20) aparece únicamente en el brazo largo de uno de los tres cromosomas 19 (flecha en figura 20).

Unificando los datos obtenidos mediante PCR y FISH se puede afirmar que se han generado y caracterizado dos líneas que mediante el uso de ZFN (#293K8 y #293K16) contienen exclusivamente una única copia KAS2.0 integrada en el genoma; más concretamente en el locus AAVS1.



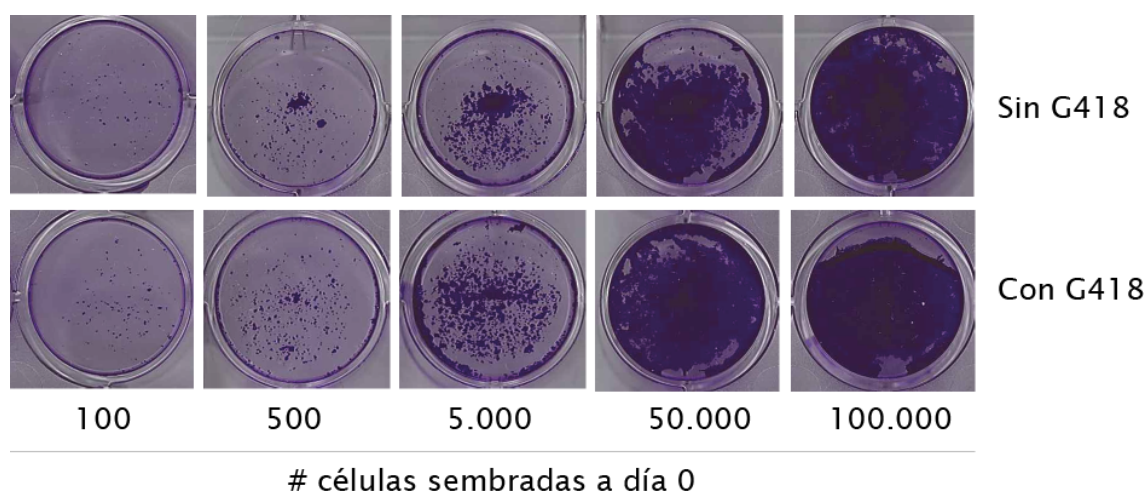
**Figura 20. Caracterización genotípica de clones 293A KAS2.0 (FISH).** Análisis por FISH de las células 293KAS2.0. (A) análisis sobre cromosomas metafásicos con la sonda genérica del cromosoma 19 (rojo, puntas de flecha) y específica del plásmido donador (verde, flecha). En los insertos se muestra a mayor aumento los tres cromosomas 19 metafásicos y los marcajes obtenidos. (B) cariotipo de uno de los clones (293K8) mostrando su hiperploidía, con tres cromosomas 19 entre otros.

### 5.1.2.2. Caracterización fenotípica de clones 293AKAS2.0

El estudio de los fenotipos de estas líneas se realizó mayoritariamente sobre el clon #293K8. Se llevaron a cabo dos ensayos para evaluar el fenotipo y la estabilidad fenotípica transferida por KAS2.0:

#### a) Ensayo de clonogenicidad

Se estudió la estabilidad del fenotipo del clon #293K8 mediante un ensayo de clonogenicidad clásico. Para ello las células se mantuvieron en cultivo durante dos meses, lo que representa más de 60 duplicaciones celulares, sin presión selectiva alguna y posteriormente se sembraron sendos cultivos con 100-100.000 células en presencia de medio DMEM normal o de medio de selección, DMEM-G. Los cultivos se mantuvieron durante dos semanas y se tiñeron para recuento de las colonias crecidas en ese tiempo. La figura 21 muestra el resultado obtenido en uno de los ensayos. El número de clones es comparable en ambas condiciones de cultivo a todas las diluciones lo que parece indicar la homogeneidad y estabilidad de este clon para el carácter neoR.

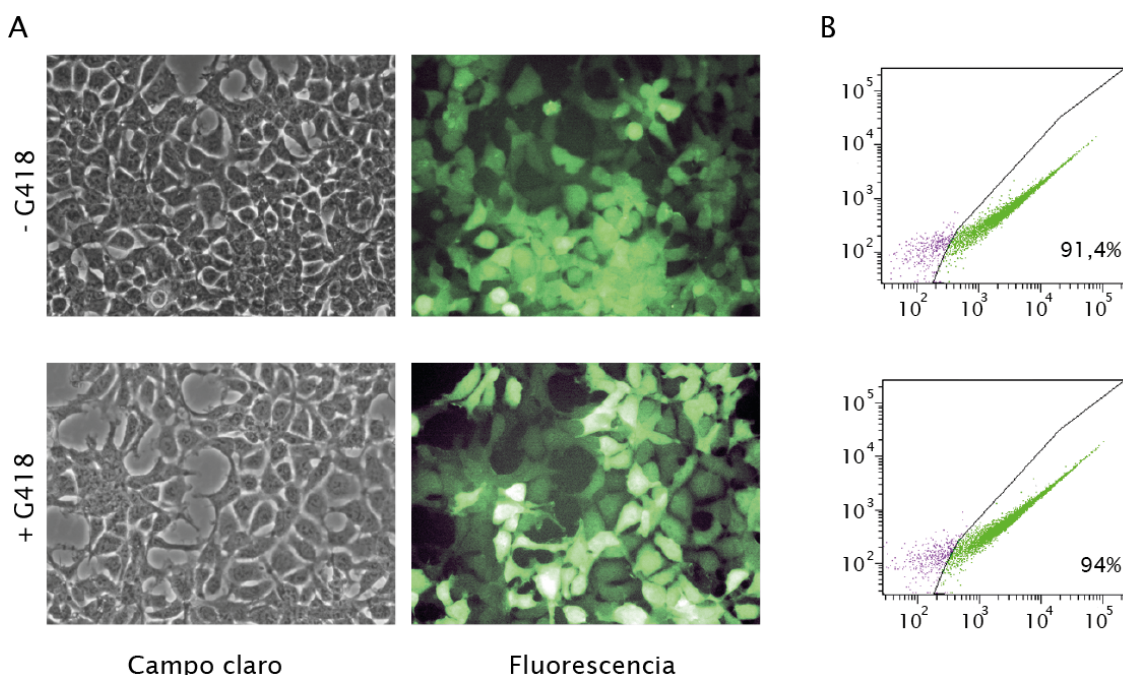


**Figura 21. Caracterización fenotípica del clon 293A K8 (I).** Ensayo de clonogenicidad del clon 293A K8 mediante recuento de colonias G418R obtenidas en ausencia (arriba) o presencia del antibiótico (abajo). Se representan placas teñidas a los 10 días de selección en cultivos sembrados con el número de células indicado.

#### b) Ensayo de estabilidad

Se estudiaron cultivos mantenidos de la misma manera que anteriormente, por análisis de FACS; para determinar la proporción e intensidad de fluorescencia (MFI) entorno a la expresión de eGFP. Asimismo los cultivos se observaron por epifluorescencia. La figura 22 resume ambos resultados, donde se aprecia que la proporción e intensidad de la expresión es casi idéntica, tanto en los cultivos mantenidos con selección, como sin ella.

Esto es indicativo de la estabilidad de la casete introducida durante cerca de dos meses en cultivo que representa al menos 60 duplicaciones poblacionales.



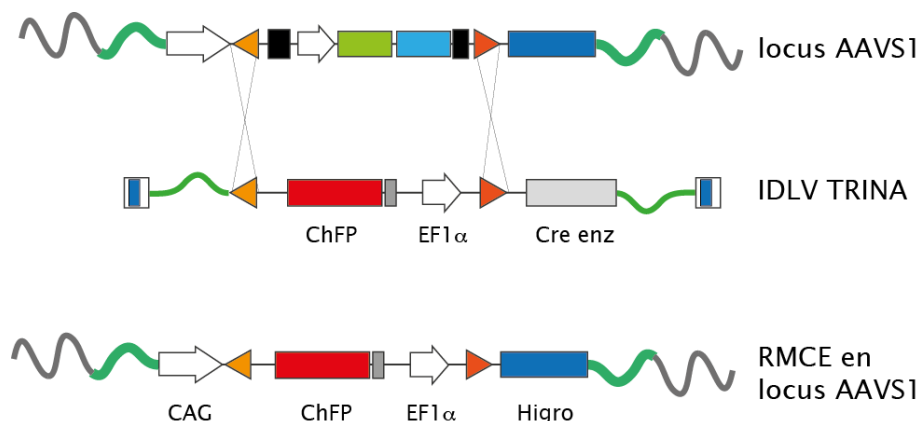
**Figura 22. Caracterización fenotípica del clon 293A K8 (II).** Ensayo de expresión de eGFP tras dos meses en cultivo con o sin selección, según se indica. (A) estudio por epifluorescencia de los cultivos y (B) análisis cuantitativo mediante FACS, se indica el porcentaje de células eGFP+.

## 5.2. Generación de clones por RMCE

Se dispone por tanto de los componentes planteados inicialmente para realizar la prueba de función sobre una única copia integrada de la casete aceptora, esto es la línea celular y los vectores IDLVs. Se diseñaron ensayos para verificar si la RMCE se produce en el locus AAVS1 y si una copia es suficiente para rescatar los fenotipos mencionados en el diseño.

Se utilizaron vectores IDLV-TRINA como vehículo que porta la casete donadora de RMCE, sobre las células #293AK8 y #293AK16 que contenían la casete aceptora y presentaban fenotipo neoR eGFP+ / ChFP- higroS. Los ensayos se monitorizarán por la ganancia de la proteína roja fluorescente y la adquisición de resistencia a higromicina, con la pérdida concomitante de la resistencia a neomicina y de la expresión de la proteína verde fluorescente (figura 23). Las transducciones se hicieron a MOI 10, para favorecer el impacto del vector lentiviral en el mayor número de células y habiendo comprobado el efecto que ésta tiene sobre la eficiencia de la RMCE (figura 14). Transcurridas 72 horas desde el momento de la transducción, las células se diluyeron 1:8 para aislar clones celulares, se inició la selección con higromicina en el quinto día post-transducción y se mantuvieron durante 10 días hasta la aparición de colonias. Veinticuatro clones se amplificaron y de ellos

se seleccionaron seis para estudios posteriores (3 procedentes de #293AK8 y otros 3 de #293AK16).



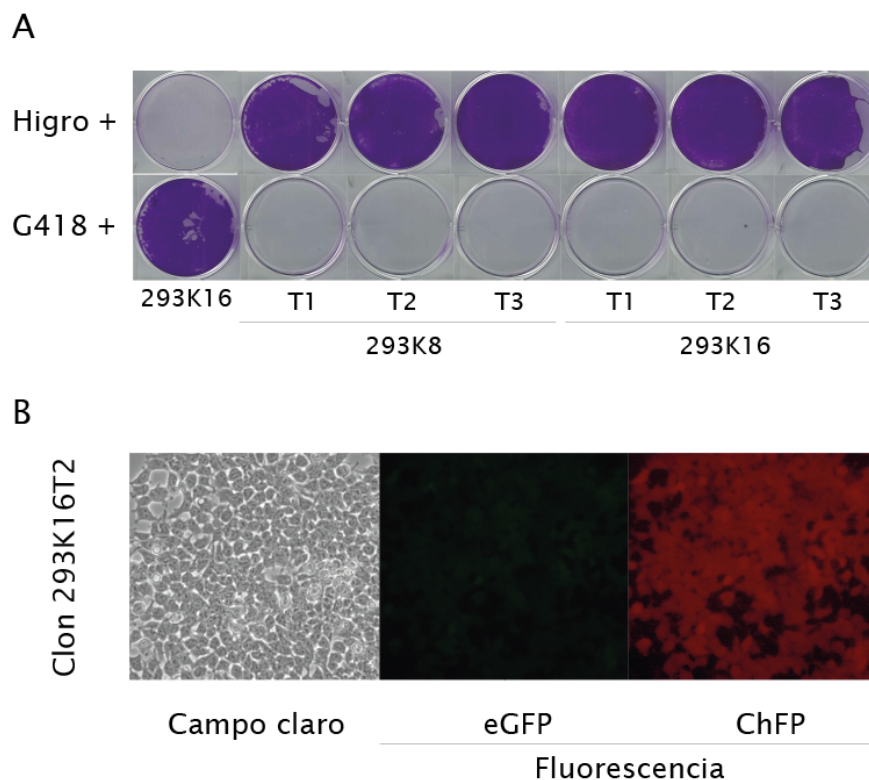
**Figura 23. Esquema ilustrativo de la reacción de RMCE** esperada entre la copia KAS insertada en AAVS1 de 293A KAS (arriba) y el cDNA del lentivirus IDLV-TRINA (centro). La reacción mediada por la enzima se simboliza por las aspas entre ambas construcciones génicas. La leyenda de los símbolos es idéntica a la empleada en figuras anteriores (6, 16 ó 19).

#### 5.2.1. Caracterización fenotípica

Los clones seleccionados de las dos líneas, #293AK8 (T1-3) y #293AK16 (T1-3), se estudiaron en su resistencia a antibióticos. Cultivos de cada uno de ellos y de los clones parentales se sembraron y mantuvieron en presencia de G418 o higromicina, tras 10 días los cultivos se tiñeron para visualizar la resistencia relativa en cada medio selectivo. La figura 24A muestra como los clones parentales son resistente a G418 y no lo son a higromicina; mientras que los 6 clones del tipo KT analizados son resistentes a higromicina y sensibles a G418, como cabía esperar del diseño experimental. Consistente con este resultado todos los clones KT mostraron fluorescencia roja (ChFP+) pero no verde (eGFP-), al contrario de lo que ocurre con el clon parental del que provienen (figura 24B).

Todo ello indica que la reacción de recombinación mediada por IDLV-TRINA permite seleccionar con antibiótico células en las que la reacción de RMCE ha sido fidedigna y con una elevada frecuencia (6 clones positivos de 6 ensayos). Queda por resolver si la selección puede realizarse en sentido inverso: aislando clones por el marcador fluorescente para obtener en frecuencia y características resultados análogos.





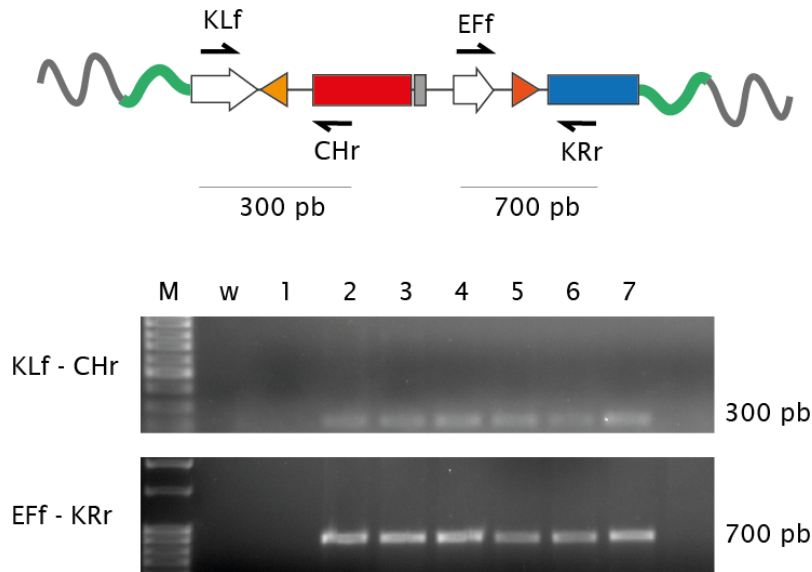
**Figura 24. Caracterización fenotípica de clones KAS2.0 aislados tras la reacción RMCE con TRINA.** (A) Estudio de resistencia a higromicina vs. neomicina en tres clones (T1, T2, T3) derivados bien de 293AK8 o de 293AK16. Igual número de células fueron sembradas en cada caso y mantenidas durante 10 días en medio DMEM-H o DMEM-G. (B) Microfotografías de uno de los clones KT, tanto de campo claro, como de fluorescencia para mostrar la conversión fenotípica tras la RMCE.

### 5.2.2. Caracterización genotípica

Se procedió al estudio genotípico de los clones #293AK8T1-3 y de #293AK16T1-3 para verificar si la reacción de RMCE ha tenido lugar única y exclusivamente en el sitio aceptor y de la manera esperada.

#### a) Análisis por PCR

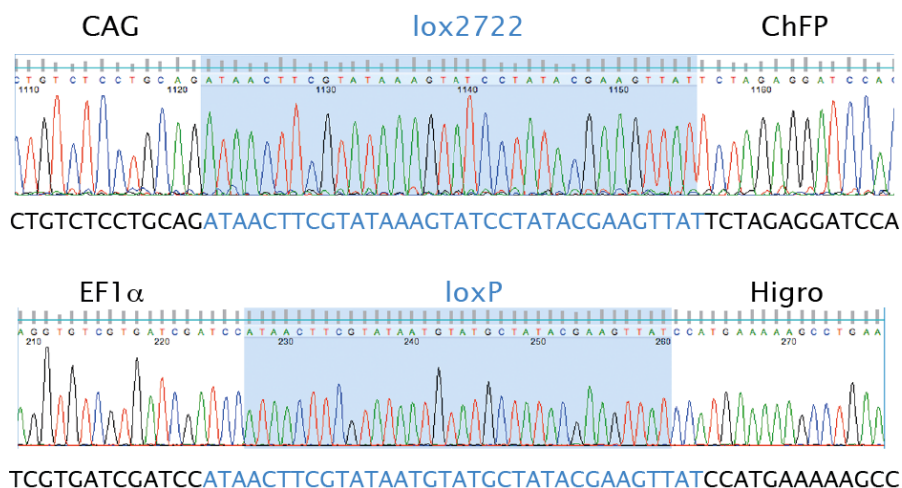
Se realizó de la misma manera que como se describió en el apartado “*Caracterización genotípica de clones 293A KAS2.0. 1.- Análisis mediante PCR*”. En este caso, los *primers* hibridan tanto en la casete de intercambio (TRINA), como en los elementos remanentes en el locus pertenecientes a KAS2.0 (región 5’ promotor CAG y región 3’ higromicina). El análisis de los dos extremos sobre el DNA genómico de los seis clones, generó las bandas esperadas de 300-pb para el extremo 5’ y la de 700-pb en el extremo 3’ (figura 25). Estos resultados son una demostración de que el intercambio RMCE ha introducido la casete donadora en el lugar esperado y por tanto apoya los cambios fenotípicos descritos en el apartado anterior.



**Figura 25. Caracterización genotípica de clones 293A KT (I).** Análisis de PCR con los *primers* indicados sobre la construcción génica esperada en los clones KT. En el panel superior se detectan productos de RMCE a través de sitio lox en 5' y en el de abajo, el correspondiente a sitios lox 3'. En ambos casos los *primers* hibridan en una región del aceptor (promotor CAG, ORF higrOR) y en una del donador (ORF de ChFP, promotor EF1a). M, marcador; w, control sin DNAg; carril 1, control 293AK8; carriles 2-4 clones 293AK8T1-3; carriles 5-7, clones 293AK16T1-3.

#### b) Análisis por secuenciación

Se quiso confirmar inequívocamente que dicha reacción había ocurrido en los sitios *loxP* y *lox2722* y que la reacción había regenerado las secuencias esperadas dejando los elementos insertados según estaba previsto. Para ello se purificaron los productos de PCR (figura 25) y se clonaron en el vector pGEM-TEasy. De las colonias obtenidas, se secuenciaron tres para cada extremo y clon. Las secuencias obtenidas se compararon con las esperadas de una reacción RMCE fidedigna. En la figura 26 se muestra el alineamiento de la secuencia para uno de los clones analizados. Se verificó que todos eran idénticos y que por lo tanto en todos ellos la reacción había procedido en la manera predicha.

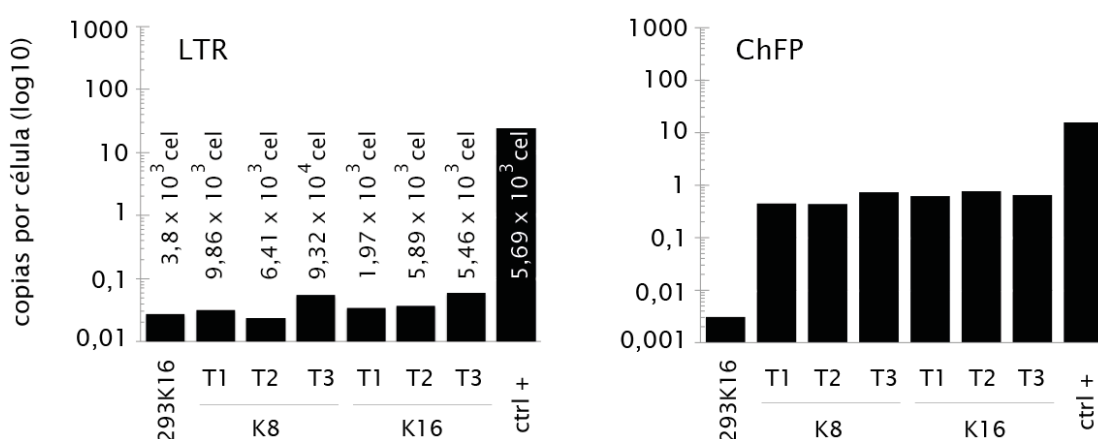


**Figura 26. Caracterización genotípica de clones 293A KT (II).** Análisis por secuenciación de los extremos implicados en la reacción de RMCE en 5' (arriba) y 3' (abajo). Se muestra la región del cromatograma correspondiente a la secuencia de los productos de PCR indicados en la figura 25, de las regiones adyacentes a los sitios lox que intervienen en la reacción RMCE. Se indican los genes *upstream* y *downstream* de dichas secuencias.

### c) Análisis mediante qPCR

Los estudios realizados confirmaban que la RMCE había tenido lugar en el locus aceptor AAVS1-KAS2.0. Sin embargo, ya que los pares de *primers* hibridan en la casete aceptora y donadora, no son descartables eventos de integración azaroso de los IDLV, en otra localización cromosómica; es por ello que se quiso comprobar mediante una técnica mas sensible dicha posibilidad. Para ello se analizó el DNA genómico por qPCR, lo cual, una vez establecidos controles adecuados y las curvas estándar de calibrado, permitirá saber el número de copias transgénicas presente en un genoma concreto.

Para ello se estableció el locus de la albúmina (cromosoma 4, GI: 178343) como gen de copia única por genoma diploide, se verifico que las células poliploides HEK293 tenían un cariotipo normal para el cromosoma 4 (donde mapea dicho locus, figura 20). Se diseñaron *primers* específicos para establecer una curva estándar que permite cuantificar el número de equivalentes celulares estudiados en cada ensayo de qPCR. Las muestras normalizadas empleando dichos *primers*, se analizaron para determinar el número de copias presentes del transgén *mCherry* en cada genoma diploide. Por último se cuantificó el número de LTRs virales en los genomas como medida de la presencia de formas integradas al azar a partir de intermedios de la retrotranscripción. Es necesario apuntar que la reacción de RMCE introduce únicamente regiones internas a las LTRs (la casete de intercambio) y no estas, aunque si se hayan contenidas en los intermedios de la retro-transcripción (cDNA viral).



**Figura 27. Caracterización genotípica de clones 293A KT (III).** (A) Análisis de la integración de las secuencias contenidas en los IDLV-TRINA por qPCR. (B) Estimación de copias de la secuencia interna de TRINA por célula (ver detalles en el texto) en los clones 293AK8/16T. Se indica el número de células empleado de cada tipo celular en el ensayo de qPCR, obtenido mediante cuantificación de copias del gen de la



albúmina (equivalentes celulares). Ctrl+: células supertransducidas con lentivirus control integrativo que expresa ChFP; 293K16, células parentales no transducidas o control negativo.

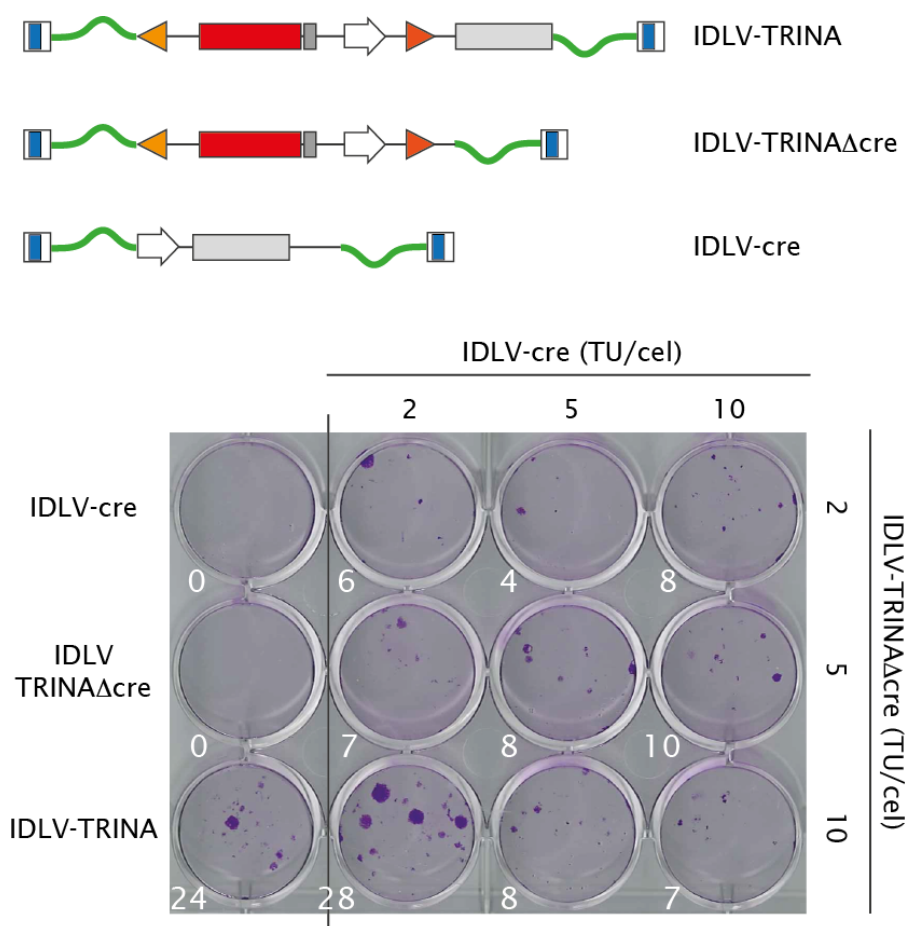
La figura 27 representa los datos obtenidos para los clones estudiados #293K8T1-3 y #293K16T1-3, lo cual indica que el número de copias del transgén recombinado es 1 por célula. Por otro lado, el número de copias detectable de LTRs es idéntico al obtenido para las células control parentales no transducidas #293K8 y #293K16 (entorno a  $10^{-2}$  copias por célula, 1 copia/100 células). Todo ello indica que el rescate de células con integraciones al azar de los IDLV es un fenómeno que, aunque con una frecuencia no nula[198], está lejos de la sinergia obtenida mediante la estrategia de *promoter trapping* y la frecuencia de RMCE.

En conjunto, los análisis de PCR, secuenciación, FISH y qPCR demuestran que en las condiciones diseñadas con el sistema KAS/TRINA se alcanzan altas frecuencias de rescate de clones; los cuales han sido recombinados por RMCE en un único sitio aceptor localizado en un locus genómico único.

### 5.3. Optimización del sistema mediante el uso de dos IDLVs

Al igual que se ha demostrado para otros sistemas de empleo de recombinasas, se evaluó cómo mejorar la entrega y eficiencia del sistema KAS/TRINA atendiendo al carácter limitante que juegan los niveles de enzima en el éxito de la reacción de RMCE (ver Resultados “Optimización del proceso de RMCE mediado por IDLV”). Para ello se generaron dos nuevos vectores (figura 28A): (i) IDLV-TRINA $\Delta$ cre que contiene la casete TRINA pero sin el ORF de *cre*; (ii) IDLV-cre que expresa solo la enzima sin casete de intercambio, por lo que su expresión no es auto-limitada, pero si transitoria (debido al carácter no integrativo de los lentivirus empleados).

Se realizó un ensayo de transducción sobre células #293K16 combinando dichos vectores y utilizando como control el IDLV-TRINA empleado en los ensayos descritos con anterioridad. Se modificó la cantidad relativa de los virus (MOI 2, 5 ó 10 TU/cel) que expresan la casete (IDLV-TRINA $\Delta$ cre) y la recombinasa (IDLV-cre), para poder valorar cuál de los dos factores compromete en mayor medida la reacción. Finalmente, como control se transdujeron cultivos con los IDLV por separado y con el control positivo (IDLV-TRINA). Tres días después de la transducción se inicio la selección por higromicina.



**Figura 28. Eficiencia relativa de RMCE mediante IDLVs alternativos a IDLV-TRINA.** (B) Colonias higroR resultantes tras la transducción de las células 293AK16 con los virus indicados en el esquema superior (A). Se indica la MOI empleada en cada condición para cada virus. En cada pocilo de la placa se muestra el recuento de colonias obtenido. Se representa un experimento de tres realizados en triplicado. El esquema mantiene los códigos de formas y colores explicado en figuras previas (en particular en la figura 23).

La figura 28 muestra un ensayo representativo, que resalta como sólo en presencia de los dos componentes (casete donadora y enzima) se pueden obtener colonias higroR: los controles transducidos únicamente con IDLV-TRINAΔcre o IDLV-cre no generaron en las condiciones ensayas clones resistentes. La mayor frecuencia de clones higroR ( $2,4 \times 10^{-4}$ ) se obtuvo con el vector IDLV-TRINA comparable a la obtenida con una ratio IDLV-cre / IDLV-TRINAΔcre de 1:5 ( $2,8 \times 10^{-4}$ ) (figura 28B). Las combinaciones en que el virus IDLV-cre se inoculó a MOI mayor a 2 (es decir, 5 y 10) pusieron de manifiesto el efecto tóxico que la enzima tenía en los cultivos, comprometiendo el resultado final. Por otro lado, cuando las cantidades de IDLV-TRINAΔcre son bajas (MOI 2 y 5) parece que la menor cantidad de sustrato de la reacción compromete la tasa de eventos de RMCE. En conjunto, la falta de sustrato y el exceso de enzima, hacen que la obtención de colonias se vea comprometida en las distintas condiciones testadas; exceptuando, el ensayo en que se consigue el compromiso entre la ratio de donador y enzima. Esto es para una MOI baja de la enzima y alta del donador (28 colonias, figura 28B).

Así, estos experimentos sugieren que si bien el sistema permite ser optimizado combinando los elementos mediante el uso de distintos virus; la opción contenida en el diseño original que contempla (i) la presencia simultánea de la enzima y la casete donadora en cada célula impactada y (ii) autolimita la expresión de la recombinasa, se acerca a una situación ideal.

## 6. Prueba de función 4. Versatilidad del sistema

### 6.1. Recombinación homóloga en línea celular primaria (Fibroblastos BJ)

Una vez caracterizado el sistema en una línea celular establecida; se decidió generar una línea similar sobre un cultivo de fibroblastos primarios humanos (BJ, ATCC CRL-2522). Al igual que en el caso de la línea celular, en los fibroblastos se testaron los distintos plásmidos necesarios para la generación de clones con copia única insertada en la localización previamente escogida (como ya se ha comentado *AAVS1*); estos plásmidos son: control, pZDKAS2.0, mRNA ZincFinger (no se trata de plásmido sino de mRNA) y mRNA ZFN junto con pZDKAS2.0. La diferencia estriba en que en esta ocasión se introdujeron los distintos componentes mediante electroporación y no por métodos basados en complejos policationicos o precipitados de cloruro de calcio, ya que ofrecen una baja tasa de transfección o una reducida viabilidad del cultivo en el tipo celular utilizado [199].

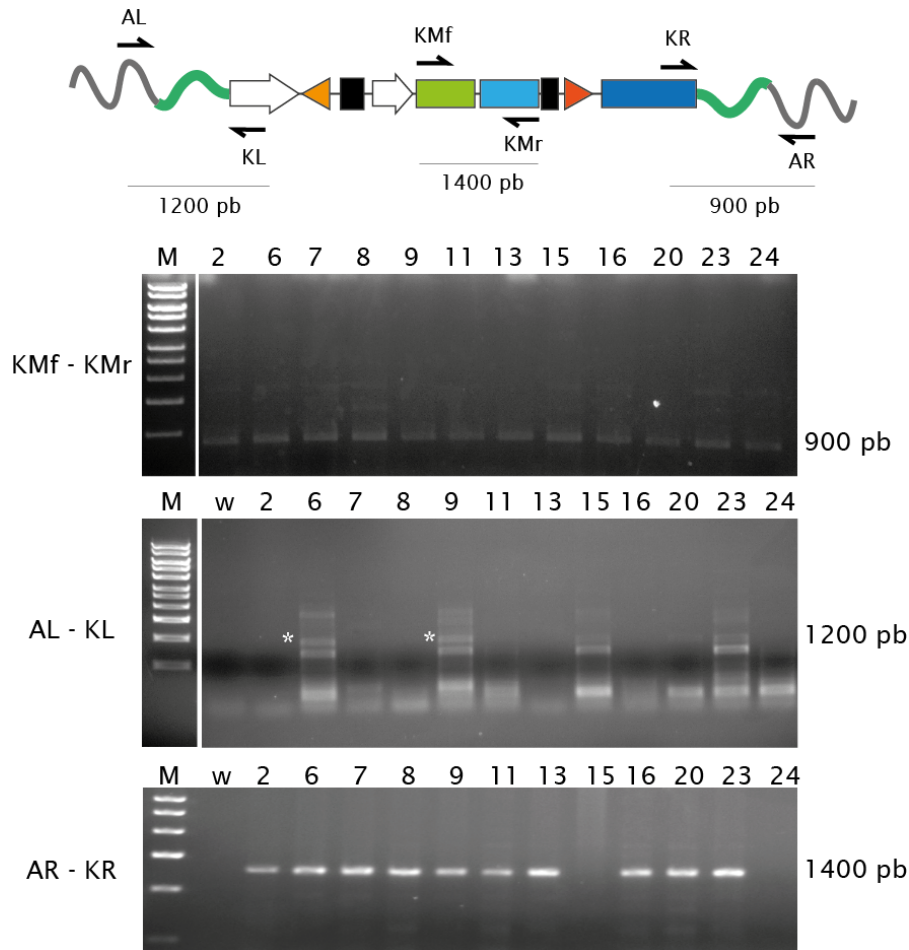
En primera instancia, fue totalmente imposible llevar a cabo la obtención de clones aislados, sino únicamente células aisladas, incapaces de proliferar clonalmente. Por ello se decidió reunificarlas en un pocillo de cultivo celular de menor diámetro; de esta manera si fueron capaces de proliferar a modo de población celular mixta (*pool*). Este hecho, unido al proceso de electroporación, afectó a su tasa de división, dado que sólo era necesario pasarlas una vez a la semana, en vez de dos como hasta el momento venía realizándose.

Dado que los datos parecían indicar que las células habían entrado en un período de senescencia o estaban muy próximas a él, junto con el hecho de que no eran clones sino un *pool* celular, llevó a la búsqueda de un método alternativo. Con tal fin se realizó una revisión bibliográfica para tratar de encontrar una solución que pudiera facilitar el aislamiento de poblaciones clonales en cultivos de fibroblastos primarios. Se optó (i) por el uso de una capa basal de fibroblastos mitóticamente inactivos (irradiados), que permitieran la clonogenicidad de las células empleadas y (ii) de un sustrato basal que estimulara el crecimiento [200-202]. Aplicando estas nuevas condiciones de cultivo, se obtuvieron más de 30 clones, de los cuales se caracterizaron 12, tanto fenotípica como genotípicamente.

#### 6.1.1. Caracterización genotípica de clones BJ KAS2.0

##### a) Análisis mediante PCR

Al igual que se realizó en apartados anteriores, la caracterización se llevó a cabo mediante PCR, diseñando parejas de *primers* que permitieran: discernir los clones celulares en los que se ha producido la inserción del casete sitio específico, de aquellos en los que esto no se ha llevado a cabo. Por tanto y como se detalló en la figura 19 se comprobó la presencia de KAS2.0, mediante el uso de *primers* internos (KMf / KMr, específicos de integración) y una combinación de *primers* externo e interno para poder detectar procesos de recombinación homóloga (ALs / KLs, en el caso del brazo de homología izquierdo y KR / ARs en el derecho).



**Figura 29. Caracterización genotípica de clones de fibroblastos BJ KAS2.0.** Se muestra el esquema de la casete integrada en el locus AAVS1 y la posición de los *primers* empleados para caracterizar tanto el número de copias como la integración en el mismo. Los paneles inferiores muestran los resultados de PCR obtenidos sobre DNAg de 12 clones G418R de los 30 aislados. La selección de *primers* en cada ensayo (izquierda) permite detectar presencia (foto superior), integración por HR en el locus AAVS1 en el brazo izquierdo (centro) o derecho (abajo) (ver texto para detalles). Los números indican el del clon aislado.

De los 12 clones que se caracterizaron 9 (clones #2, #6, #7, #8, #9, #11, #16, #20 y #23) fueron capaces de proliferar con normalidad; mientras que 3 (clones 13, 15 y 24), entraron en senescencia y no sobrevivieron al proceso de expansión clonal, aunque se logró obtener DNA para su análisis inicial. De todos ellos, únicamente dos (clones #6 y #9) fueron

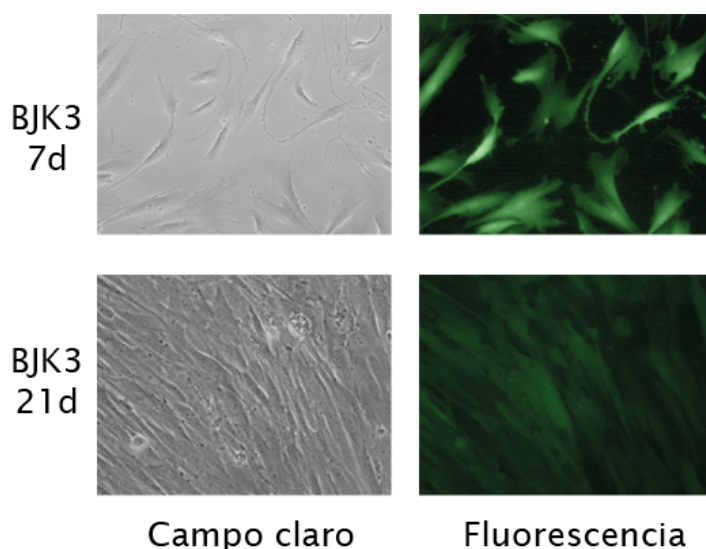
positivos tanto para el brazo izquierdo de homología, como para el derecho (figura 29). Ello significa una frecuencia de HR mediado por la ZFN, en estas células, de 20% sobre los supervivientes, o del 16% sobre los 12 estudiados. Por tanto se disponía de dos clones sobre los que poder llevar a cabo los subsiguientes ensayos.

#### 6.1.2. Caracterización fenotípica de clones BJ KAS2.0

Para llevar a cabo dicha caracterización, uno de los clones (BJK6) fue el objeto experimentación.

##### a) Ensayo de estabilidad

Tras el período de selección y expansión clonal, se retiró el agente selectivo a las células de tal manera que carecieran de presión selectiva y por tanto se pudiera caracterizar la estabilidad del transgén o casete introducido. Tras 3 semanas en cultivo, dichas células se observaron por epifluorescencia y mostraban un alto porcentaje con la fluorescencia verde mantenida (figura 30), lo que confirma que el casete introducido es estable a lo largo del tiempo.



**Figura 30. Caracterización fenotípica del clon BJK6.** Expresión de eGFP tras 3 semanas en cultivo del clon BJK6 en ausencia de selección con G418. Se muestran microfotografías tanto de campo claro, como de epifluorescencia.

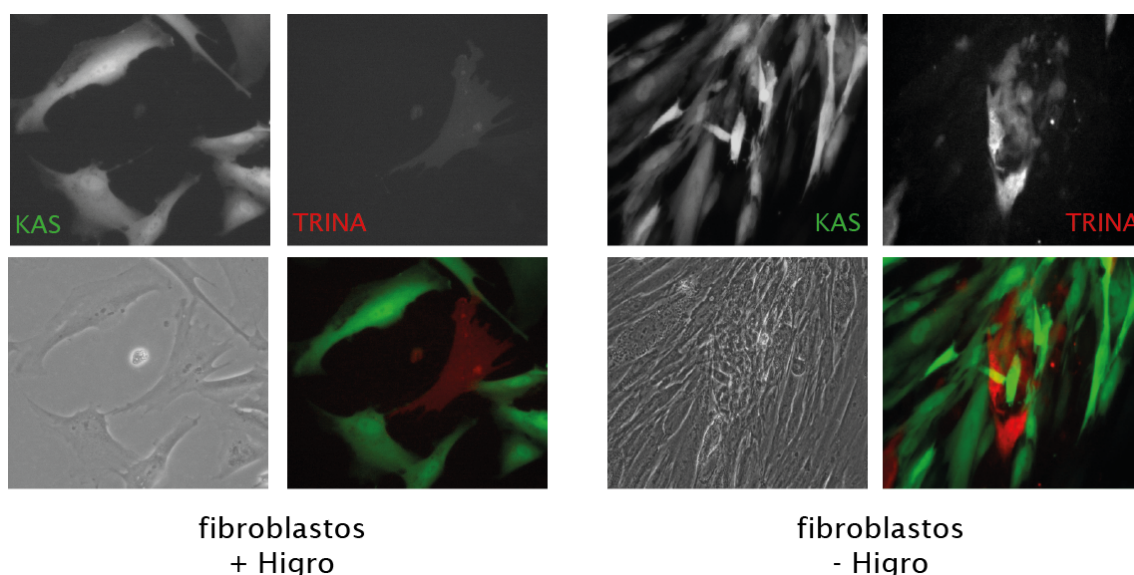
#### 6.2. Generación de clones BJ por RMCE

Una vez se disponía de la línea primaria sobre la que llevar a cabo los experimentos de RMCE y dicha línea es estable, se probará la recombinación sobre ella. Se transdujeron a

MOI 5 células BJK6 tanto con un IDLV-TRINA, como con el control IDLV-CMV.eGFP. Transcurridas 72 horas, desde el momento de la transducción, las células se diluyeron 1:8 para aislar clones celulares, se inició la selección con higromicina en el quinto día post-transducción y se mantuvieron durante 10 días. Al igual que había sucedido anteriormente con la generación de la línea parental BJK6, se observaron problemas en la obtención de clones; pero en este caso, no pudieron solventarse variando las condiciones de cultivo como anteriormente. Por ello se decidió continuar el experimento con el cultivo en forma de población mixta (*pool* de células transducidas).

### 6.2.1. Caracterización fenotípica

Dado que la población aislada BJK6Tp (TRINA *pool*) no era capaz de crecer de una manera clonal, se estudió el cambio fenotípico en el marcaje fluorescente (pérdida de eGFP y adquisición de ChFP). La figura 31 muestra que parte de las células obtenidas han cambiado el marcador eGFP, el cual se pierde, concomitante con la aparición de células ChFP+.



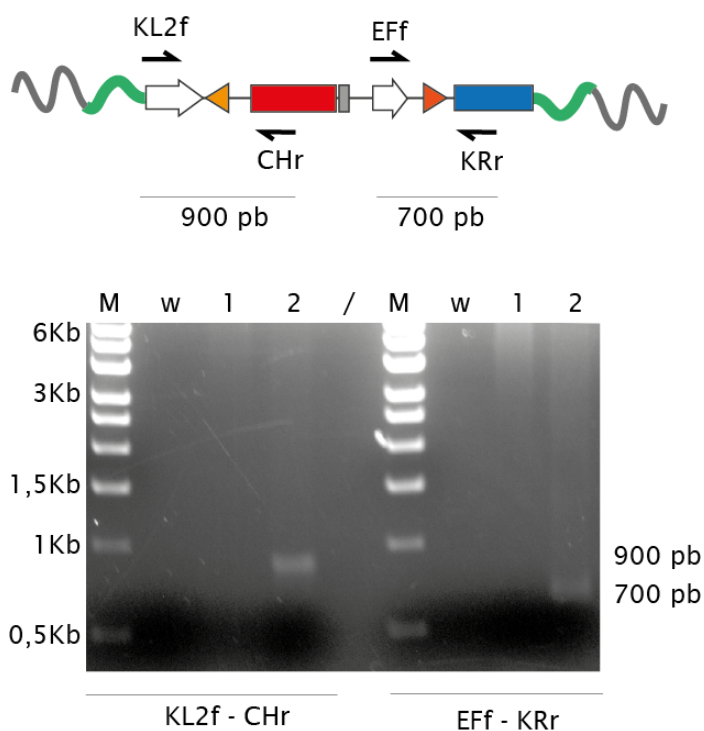
**Figura 31. Caracterización fenotípica del pool BJK6Tp tras transducción con IDLV-TRINA.** La reacción de RMCE se puso de manifiesto mediante la aparición de células ChFP+, sobre un cultivo eGFP+ (BJK6). Se muestran fotografías de fluorescencia de un campo representativo de los cultivos en presencia o ausencia de higromicina para la detección de eGFP (KAS) o ChFP (TRINA), y la fotografía coloreada *merge* de las anteriores.

Se puede apreciar que debido a la incapacidad de aislar clones, e incluso de llevar a cabo la selección por antibiótico hasta último término (pues aunque se morían las células KAS, se veía impedida la proliferación de células TRINA, y eran difícilmente rescatables), lo que se dispone es de una población mixta de BJK6 y BJK6Tp; eso sí, las imágenes parecen indicar que las células ChFP no eran eGFP y viceversa.

### 6.2.2. Caracterización genotípica

Como no se aislaron clones y la presencia del agente selectivo era deletérea para las células, se amplificaron sin él. En dicha población mixta, se procedió al estudio genotípico caracterizando el emplazamiento en el que se ha introducido la casete donadora, es decir, completar la verificación de si la reacción de RMCE había tenido lugar única y exclusivamente en el sitio aceptor y de la manera esperada.

Al igual que en los apartados anteriores (*Análisis mediante PCR*), se estudió mediante el empleo de parejas de *primers* combinando externos e internos a la zona de recombinación, para poder dilucidar si el evento se llevó a cabo en el emplazamiento previsto. El análisis de los dos extremos sobre el DNA genómico del *pool* celular (figura 32), generó las bandas esperadas de 900-pb para el extremo 5' y la de 700-pb en el extremo 3' (ver esquema en dicha figura). Estos resultados son una demostración de que el intercambio RMCE ha introducido la casete donadora en el lugar esperado y por tanto apoya los cambios fenotípicos descritos en el apartado anterior.



**Figura 32. Caracterización genotípica del pool celular BJK6Tp.** Análisis de PCR con los *primers* indicados sobre la construcción génica esperada en el pool de células BJK6 transducidas con IDLV-TRINA. La aparición de las bandas de tamaño indicado representa la presencia de células en el cultivo en las que ha tenido lugar el intercambio. (M) marcador MW; (w) control sin DNA; (/) carril vacío; (1) control BJK6; (2) BJK6Tp.

Por tanto, se pueden concluir varios aspectos, (i) la reacción de RMCE es independiente del tipo celular utilizado, (ii) se ha de tener en cuenta tanto la capacidad clonal del tipo



celular de trabajo, como (iii) el efecto de emplear agentes selectivos. Estos factores en su conjunto, permitirán que la estrategia utilizada tenga una mayor probabilidad de éxito o la obtención de los resultados perseguidos.



## MODELO 2

Sistema de entrega de enzimas de recombinación

Aplicación a la integración sitio específica mediada por rep78 de AAV para inserción en el locus humano *AAVS1*.

### 1. Objetivo del diseño

La modificación del genoma humano con fines de investigación o en estrategias terapéuticas por integración dirigida de un transgén (como compensación génica, no de reemplazamiento génico), es uno de los principales problemas que existe en el avance y/o aplicación para terapia génica [36]. Dicha contingencia es el resultado de la falta de sistemas seguros y eficientes que puedan realizar la entrega de material génico sitio específico o que aún poseyendo cierta integración específica, como ocurre con algunos sistemas virales, carezcan de (i) una expresión estable dependiente de locus, (ii) de una alta capacidad de carga (transportar material génico) y (iii) falta de universalidad, es decir, que sean aplicables a un amplio rango de tipos celulares, tanto *in vitro* como en su posible extrapolación a sistemas *in vivo*. Por todo ello, se definen una serie de requerimientos mínimos para que los sistemas de entrega génica sean óptimos:

- a) La frecuencia de eventos de integración sitio específica vs. al azar debe estar maximizada hacia la primera para de esta manera facilitar su rescate y tener una mayor aplicación.
- b) Dado que dichas integraciones están mediadas por la actividad de distintas enzimas aportadas en *trans*, según cual sea el sistema escogido se debe disponer de una manera eficiente y controlable de hacer su entrega al interior celular. Además al ser citotóxicas y/o deletéreas para la célula (debido a la generación de cortes en una o las dos hebras del DNA) requieren de una regulación muy fina en su expresión y temporalidad.
- c) Permitir diferenciar los eventos de integración aleatorios de aquellos que se sucedan en el sitio de elección o genómicamente seguro, mediante la utilización de un sistema reportero adecuado.
- d) Interferir mínimamente en la homeostasis celular, minimizando el empleo de herramientas genéticas/moleculares exógenas, pero sin comprometer las capacidades de introducir el transgén en una gran variedad de tipos celulares de manera segura y eficiente.

El objetivo es complementar con nuevas funcionalidades, que pudieran contribuir a su modificación favoreciendo la integración sitio específica mediante una aproximación alternativa al método de marcaje y reemplazamiento previamente empleado (resultados “Modelo #1”). Con este fin se quiere hacer uso de los elementos mínimos que de manera natural confieren integración sitio específica a los virus adeno-asociados (AAV), en un proceso de infección latente [54]. Para ello se pretende combinar el uso de la proteína rep78/68 y una secuencia de reconocimiento mínima.

## **2. Elementos del diseño**

### **2.1. Fusiones proteicas**

Una manera altamente eficiente de introducir proteínas biológicamente activas en células eucariotas es la transducción proteica [161]. Como se muestra en la figura 33, las partículas han de contemplarse como verdaderos reservorios proteicos ya que hay una gran variedad y cantidad de proteínas encapsidadas en el interior de una partícula

En definitiva, se generará una herramienta de entrega de alto número de proteínas

## 2.2. Proteínas testadas

Se desarrollará el sistema de entrega para distintas proteínas, con el objetivo de poder caracterizar, como fin último, la integración sitio específica mediada por una de ellas (rep78). Por ello, se ha seguido el siguiente esquema que avanza secuencialmente en complejidad y prueba de función de la viabilidad de la aproximación: eGFP – cre – rep78 cada uno de ellos aportará respuesta a alguna de las preguntas que se plantean.

- a. eGFP: ofrecerá respuesta a las preguntas de si es posible la encapsidación, visualización y su potencial transductor (entrega).
- b. Cre: la recombinasa permitiría responder a la pregunta de si pueden entregarse proteínas funcionales con la aproximación planteada.
- c. Rep78/68: usar este  
sistema como un método de integración de DNA sitio específico.

## 3. Estabilidad y producción de la partículas

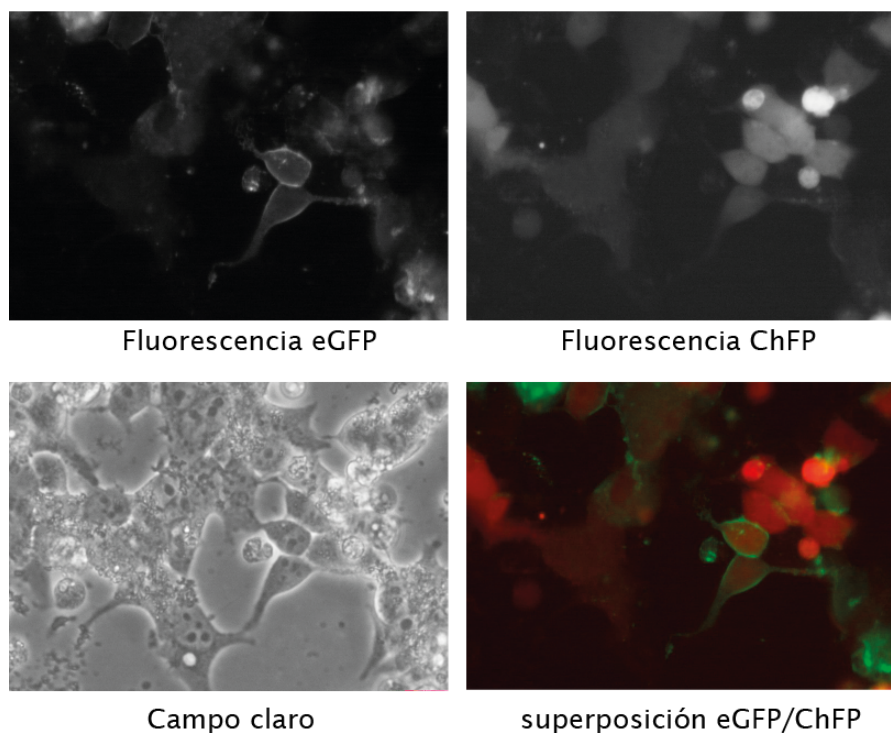
### eGFP

### 1.1. Producción y titulación de particulares

### de entrega de eGFP

El gen eGFP se clonó en dicha posición en el plásmido final  
y se empleó junto con el plásmido que  
expresa bajo un promotor ubicuo y fuerte (CMV) la proteína mCherry.

## células 293T



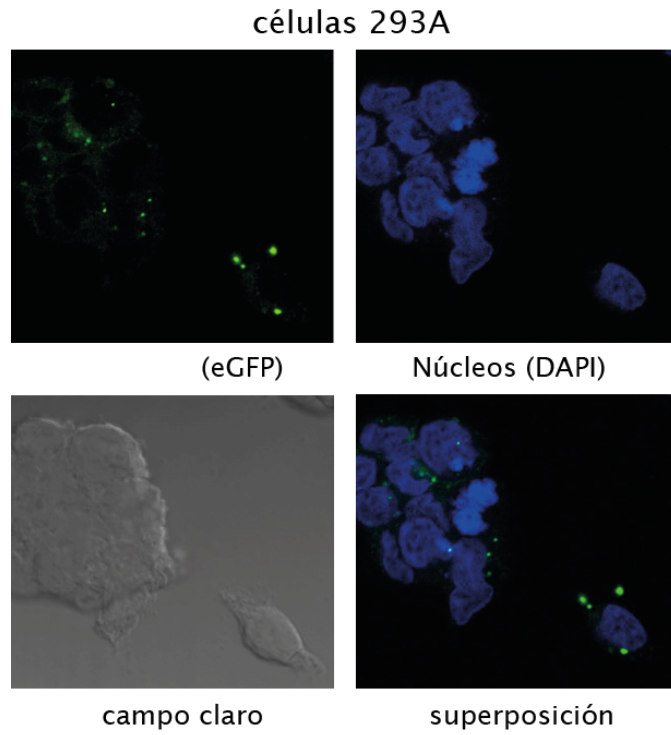
**Figura 34. Expresión de la proteína eGFP desde** . Se muestran microfotografías de epifluorescencia de células HEK293T transfectadas . La señal ChFP corresponde a células expresando el marcador del plásmido , y la expresión de eGFP a la de

La titulación por qPCR de indicó que la producción no se veía afectada significativamente por la inclusión de eGFP , al ser los títulos obtenidos ; los mismos títulos que rutinariamente se obtienen en el proceso de generación .

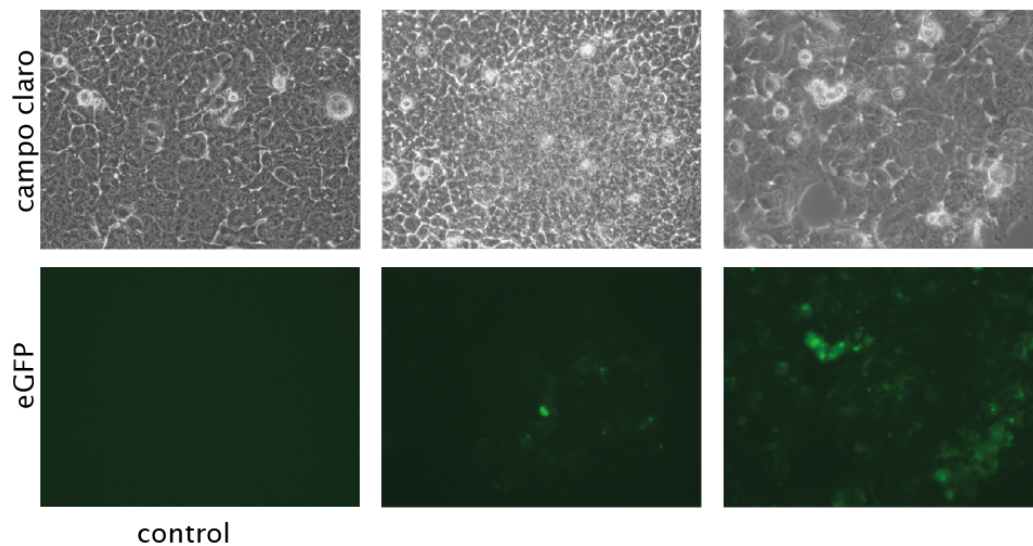
### 3.2. Transducción con partículas eGFP

Para analizar si dichas partículas transductoras de proteínas eran funcionales, se diseñó un ensayo en el que células HEK293A se incubaron con el concentrado

Los cultivos fueron fotografiados bajo epifluorescencia a distintos tiempos post-transducción (4, 16, 24, 48, 72 y 96 horas) y se fijaron para su análisis por microscopía. En los cultivos transducidos se puede apreciar el punteado debido a en las membranas celulares (figura 35), ausente en los controles sin la fusión con la proteína fluorescente. Además, se pudo apreciar que la proporción de focos eGFP+ que aparecían titulaban con la (figura 36), lo que indica una relación directa en la presencia de proteína en el concentrado .

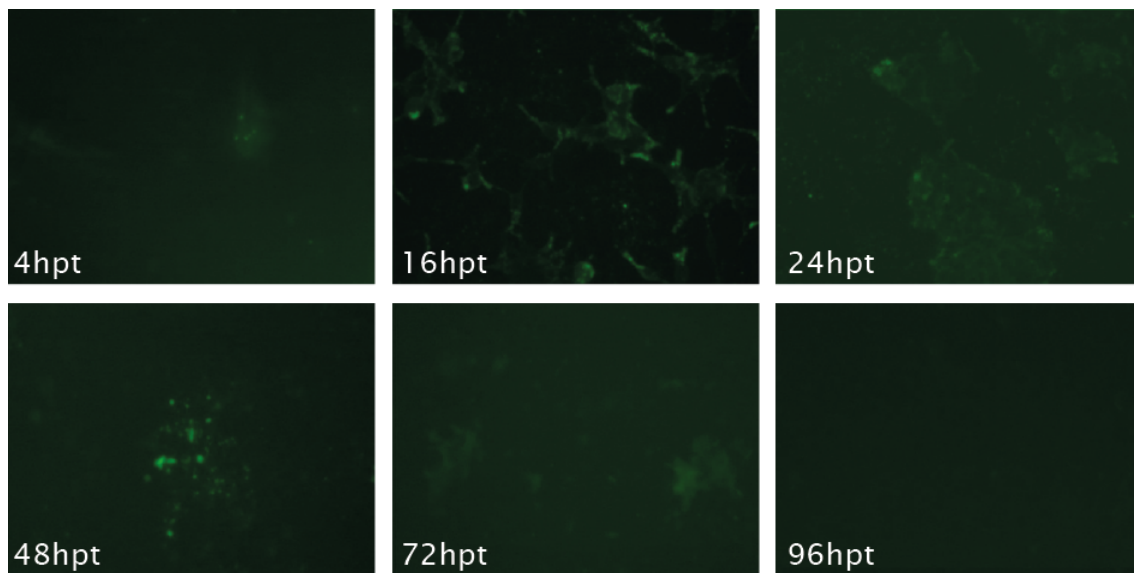


**Figura 35. Entrega de la proteína eGFP mediante transducción** . Se representan la presencia de señal de eGFP en células transducidas y fotografiadas a las 24 hpt y reveladas mediante epifluorescencia.



**Figura 36. Entrega de eGFP por** . Se muestra la señal de eGFP en células transducidas y reveladas a las 48 hpt.

Por otro lado, la cinética en la expresión e intensidad de la señal fluorescente (figura 37) fue un hecho claro: comienza a las 4hpt y aumenta hasta alcanzar un pico a las 16-48hpt conservando la distribución punteada presente en la membrana de las células impactadas, aunque con cierta difusión citoplasmática de la fluorescencia (figura 37 24hpt); a las 72 hpt la señal no es predominante en la membrana, sino que se produce un patrón citoplasmático ubicuo y difuso (figura 37 72hpt), con ausencia del punteado inicial. La señal deja de observarse transcurridas 96 horas (figura 37 96hpt).



**Figura 37. Cinética y distribución subcelular de eGFP entregada** de fluorescencia en células HEK293A transducidas a los tiempos indicados. . Se muestra la señal y fotografiadas

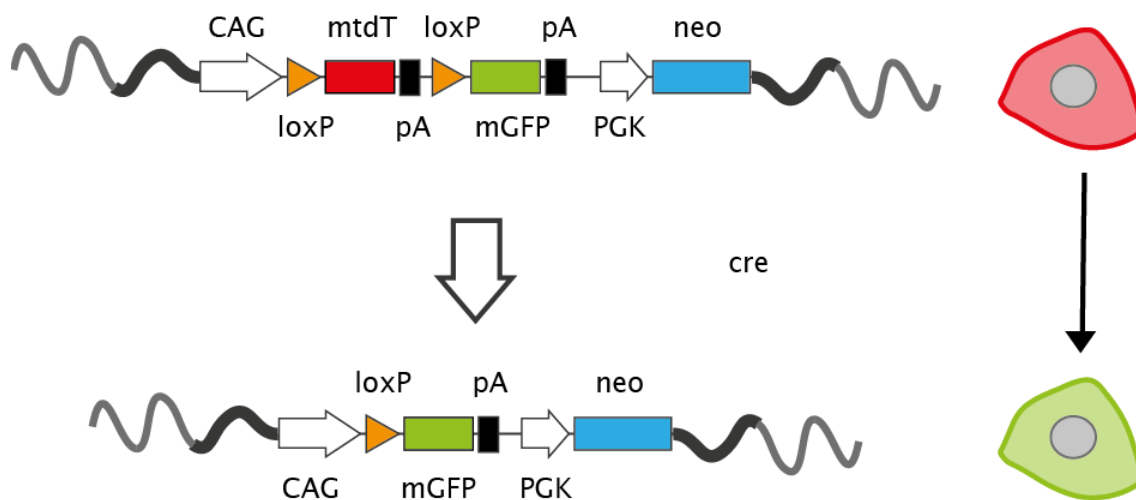
Parece por tanto inequívoco que el sistema de entrega proteica mediado es funcional y tanto su temporalidad, como su distribución espacial, es dependiente de la vida media de la proteína y el proceso de internalización .

## 4. Funcionalidad de la recombinasa cre

### 4.1. Generación y verificación de un sistema reportero

Una vez se alcanzó el primer objetivo, la entrega de proteínas se trató de dar respuesta a la siguiente pregunta: ¿podrían las proteínas entregadas ser funcionales y retener su actividad biológica intacta?. Se evaluó mediante el empleo de la recombinasa cre, como modelo de estudio de enzima que interfiere sobre el genoma. En primer lugar se generó una línea celular que contuviera integrado de manera estable un sistema reportero entre sitios loxP (idénticos); de tal manera que la acción de la recombinasa produjera la reacción de escisión y conllevara un cambio en el fenotipo celular fácilmente discernible (figura 38).

Con dicho fin, se tomó como base el sistema previamente descrito por Liqun Luo y colaboradores [206], el cual hace uso de un doble sistema reportero. Se trata de un sistema en el que se dirige la expresión del marcador de membrana rojo (*membrane tdTomato* o mT) bajo la acción del promotor fuerte CAG. Dicho marcador está constitutivamente expresado y el ORF se encuentra enmarcado entre dos sitios de recombinación loxP. Una vez que la recombinasa actúa, dicho marcador es escindido y el promotor pasa a dirigir la expresión de una versión de eGFP que localiza en membrana (*membrane eGFP* o mG). Dicha línea celular se generó por transfección y posterior selección con antibiótico (G418). La línea generada se denominó HEK293A mT/mG.

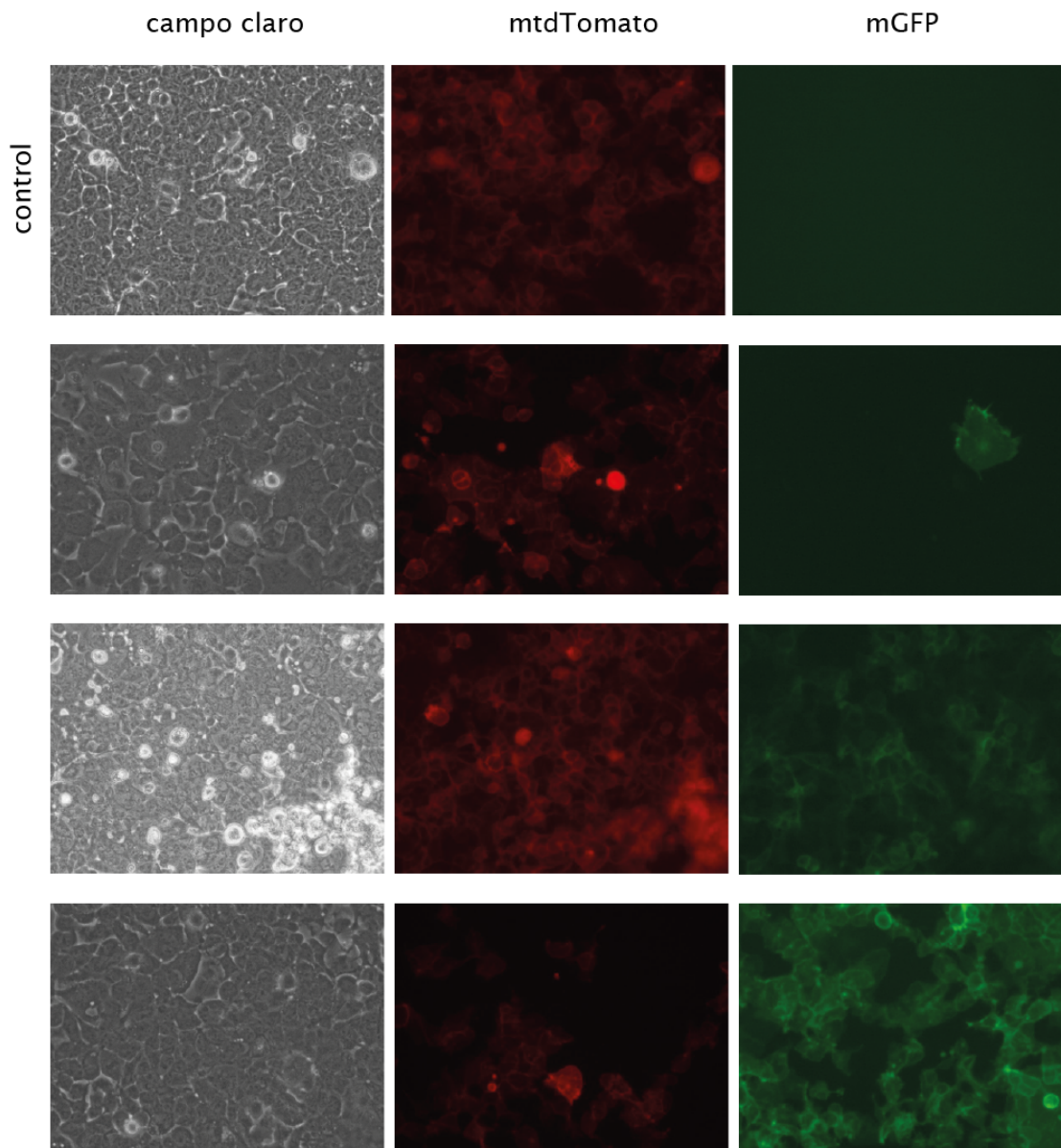


**Figura 38.** Esquema del sistema reportero empleado para caracterizar la actividad cre entregada. Los elementos promotor, ORF y lox se representan como en figuras anteriores.

#### 4.2. Estudio funcional de la proteína cre transducida

Para estudiar si dichas las partículas transductoras eran funcionales se transdujeron las células generadas y se analizaron a distintos tiempos por epifluorescencia; para de esta manera poder hacer un seguimiento en el cambio de fenotipo mtdT a eGFP. Los resultados obtenidos (figura 39) indican que al igual que sucedía con la versión , la actividad cre titula y si se analizan los cultivos, mediante FACS, una semana después de la transducción (figura 40), se aprecia la existencia de una relación directa entre ambos parámetros y ratio mtdT/eGFP).

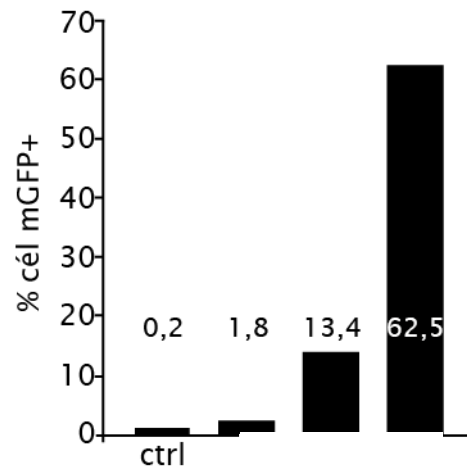




**Figura 39. Cambio fenotípico promovido por cre.** Se muestran fotografías de epifluorescencia de cultivos de HEK293mT/mG transducidos empleando cre. La funcionalidad de cre se manifiesta por la aparición de células verdes (ver figura 38).

En conjunto, estos datos nos indican que la recombinasa cre de la misma manera que lo sería la proteína original [206]. Además, dado que la relación es directa, la dosis puede ser controlada, lo cual es un factor importante en la utilización de enzimas deleteras para la célula, como se comentó en el apartado “Objetivo del diseño”.





**Figura 40. Cuantificación de la actividad cre entregada** . Se muestra la cuantificación obtenida por FACS de la proporción de células eGFP+ en los cultivos mostrados en la figura 39.

## 5. Aplicación a la integración dirigida en AAVS1.

Como se ha comentado con anterioridad, a la hora de mediar la integración dirigida sitio específica, no sólo es necesaria la enzima que produce dicha recombinación, sino que se necesitan zonas y/o secuencias de homología en las que llevar a cabo el proceso. La integración del genoma de AAV en el ciclo latente requiere de algunos elementos imprescindibles (ver Introducción apartado “integración natural en sitio específica. Virus AAV”) de entre los cuales se ha dilucidado los elementos mínimos: actividad *trans* de rep78, y secuencia en *cis* . Dichos elementos se van a introducir para estudiar la posible integración en AAVS1.

### 5.1. Generación de un sistema reportero

Para analizar dicho proceso y caracterizar la eficiencia, se diseñó un sistema reportero que permitiera diferenciar entre (i) los eventos mediados por integración al azar y (ii) aquellos que se llevan a cabo por medio de la replicasa.

El sistema reportero

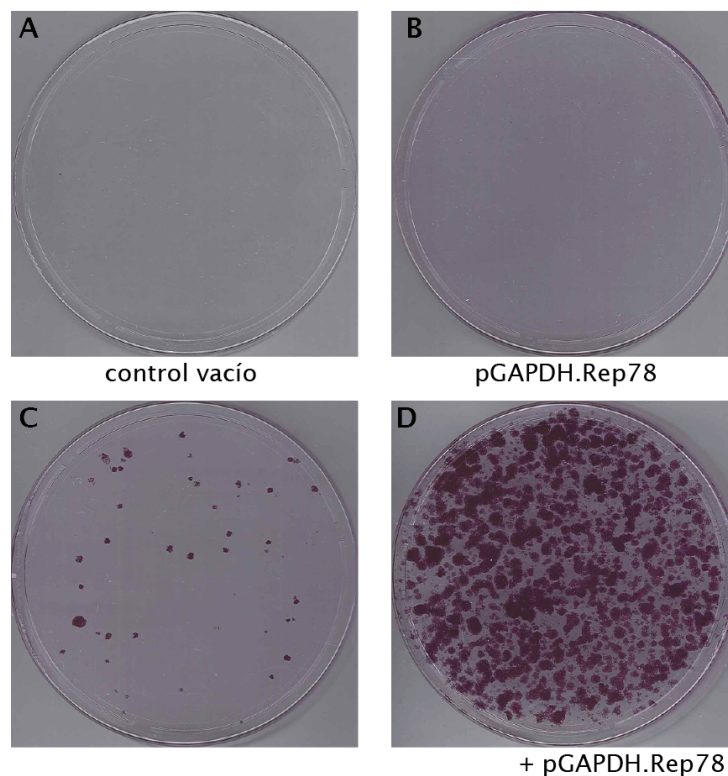
se diseñó con el fin de permitir caracterizar de manera inequívoca aquellas células en las que se hubiera producido la integración sitio específica.



## 5.2. Prueba de función

### (I). Transfección transitoria con plásmidos

La funcionalidad del sistema se testó inicialmente mediante el empleo de plásmidos transfectados transitoriamente, siendo capaces de visualizar células positivas y de aislar clones resistentes. Se utilizó el plásmido reportero y un plásmido (pGAPDH.Rep78) que expresa la proteína rep78 bajo el promotor ubicuo GAPDH. Se transfectaron cultivos de HEK293A con ambos plásmidos o con cada uno de ellos independientemente. Transcurridos 5 días los cultivos se diluyeron 1:6 en placas de 10cm<sup>2</sup> y se mantuvieron en medio de selección (DMEM-G a 450 ug/mL) durante 14 días, período en el que no quedaron células vivas en la placa control sin transfectar (figura 42A). Tras dicho período se fijaron y tiñeron para su posterior cuantificación. Sólo los cultivos que tenían el sistema reportero contenían clones resistentes y no el cultivo con la enzima o, como se ha comentado, con el control vacío. El número de clones resistentes era muy superior en el caso de la co-transfección, sobre el que únicamente poseía el sistema reportero (figura 42D y 42C, respectivamente), lo que parece indicar que la presencia de ambos componentes tiene un efecto potenciador del reordenamiento de la casete y rescate del fenotipo esperado.



**Figura 42. Análisis de la funcionalidad del sistema reportero**. Se muestran las colonias G418R obtenidas en cultivos transfectados con los plásmidos que expresan sólo rep78 (B); el sistema reportero (C) o con ambos (D). Los cultivos se mantuvieron durante 14 días en medio de selección antes de teñirse.

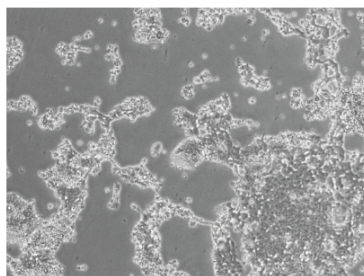
### 5.3. Prueba de función rep /

### (II): Transducción con partículas

Debido a los prometedores resultados obtenidos, se decidió estudiar la aproximación como sistema de entrega de los componentes. Al igual que en los casos anteriores

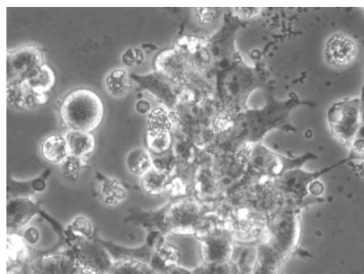
concentrados fueron empleados para transducir cultivos de HEK293A. Una primera observación mostró la citotoxicidad asociada a la enzima, ya que los cultivos transducidos conteniendo rep78 ( ) tras 24hpt sufrieron un proceso de estrés que posteriormente afectó significativamente a la viabilidad del cultivo (figura 43A) especialmente en aquellos casos que habían sido transducidos a una alta ; mientras que en el resto de condiciones, la viabilidad no se vio comprometida.

A



4x

B



40x

**Figura 43. Entrega de rep78 desde vectores** . (A) La transducción de células HEK293A con rep78 tiene un efecto sobre la viabilidad celular evidente al microscopio de contraste de fases.

A los 7dpt se analizaron los cultivos por citometría de flujo a excepción de los transducidos , como se comentó anteriormente. Los datos obtenidos (figura 43B), no mostraron diferencia alguna entre las distintas condiciones ensayadas, por lo que se pudo concluir que la estrategia contenía elementos que no permitían generar los resultados previstos a diferencia de lo observado por transfección. Todo ello llevó al planteamiento de una estrategia diferente, dado que esta resultaba inviable en la consecución de los objetivos.

#### **5.4. Prueba de función (y III):**

Con este protocolo se realizó un ensayo en el que se transdujeron cultivos de células HEK293A con cantidades variables

Se observó, en primer lugar, que en todos los casos los cultivos tenían un estado normal y su viabilidad no se veía alterada con respecto al control

). Por ello, 72hpt se sub-cultivaron en placas de 10cm<sup>2</sup>, en medio de selección DMEM-G, y tras 14 días se fijaron y tiñeron las colonias para su análisis. El número de colonias obtenidas era entre 10 y 50 veces superior en las situaciones experimentales planteadas con respecto a los controles; (figura 44, p.ej. 50 y 382 vs. 0 y 31 colonias, respectivamente).

|                    |      | (part./cel) |          |          |          |
|--------------------|------|-------------|----------|----------|----------|
|                    |      | 0,75        | 3        | 12       |          |
| rep<br>(part./cel) | 0,2  | 130 (4)     | 222 (15) | 350 (60) | Ensayo 1 |
|                    | 0,75 | 50 (1)      | 148 (4)  | 224 (16) |          |
|                    | 3    | 51 (0,25)   | 137 (1)  | 134 (4)  |          |

|                    |      | (part./cel) |         |          |          |
|--------------------|------|-------------|---------|----------|----------|
|                    |      | 0,75        | 3       | 12       |          |
| rep<br>(part./cel) | 0,2  | 50 (4)      | 60 (15) | 290 (60) | Ensayo 2 |
|                    | 0,75 | 38 (1)      | 194 (4) | 382 (16) |          |
|                    | 3    | 52 (0,25)   | 168 (1) | 364 (4)  |          |

**Figura 44. Efecto de la temporalidad y de las cantidades relativas en el rescate de clones resistentes.** Se muestra la cantidad absoluta de colonias G418R obtenidas a los 14 días de la transducción, tanto para el proceso de transducción simultánea (**A**), como para la transducción seriada (**B**). Entre paréntesis se expresa la ratio que existe

Como se puede observar en la figura 44, tras la transducción concomitante (figura 44A), el mayor número de colonias (350) se obtiene , siendo la relación 60:1 con respecto proteína rep78. Sin embargo, con niveles crecientes de la enzima, disminuye el número de colonias obtenida (214 a ratio 16:1 y 134 a ratio 4:1) reflejando el posible efecto deletéreo que rep78 tiene en la viabilidad celular. Por otro lado, cuando se comparan los resultados a ratios constantes, el efecto es idéntico, con independencia de las cantidades absolutas ; obteniéndose el mismo número de colonias tanto a ratio 16:1 como 4:1.

En todos los casos al aumentar -rep, es menor el número absoluto de colonias. Ello sugiere que el número de sustratos es determinante y que la enzima está en exceso en todos ellos.

Cuando la transducción se realizó dejando transcurrir      horas entre la introducción de los      y la enzima (      rep), se observó que al aumentar las ratios (igual que en el ensayo anterior, 1:1, 4:1 y 16:1) la eficacia del proceso, medido como número de colonias G418<sup>R</sup>, es mayor (figura 44B). El incremento del número de colonias se obtiene con el aumento del      sustrato      ), que parece por tanto de nuevo el factor limitante de la reacción; ya que en todos los casos, al incrementar la cantidad de la enzima      no se incrementa el número de colonias de manera notoria. Cabe destacar pues, que la previsible disponibilidad de moldes/sustrato,      , en el momento en que se hace disponible la enzima, es un factor que facilita la reacción, y por tanto la eficacia con que se obtienen colonias.

Con todo ello, se puede concluir que se ha desarrollado un sistema reportero funcional de integración dependiente de rep78, que permite rescatar clones con una eficiencia dependiente tanto de

Queda por demostrar si y en que frecuencia dichos clones han sido integrados en el sitio AAVS1 como factor que se anticipó necesario para el rescate fenotípico.

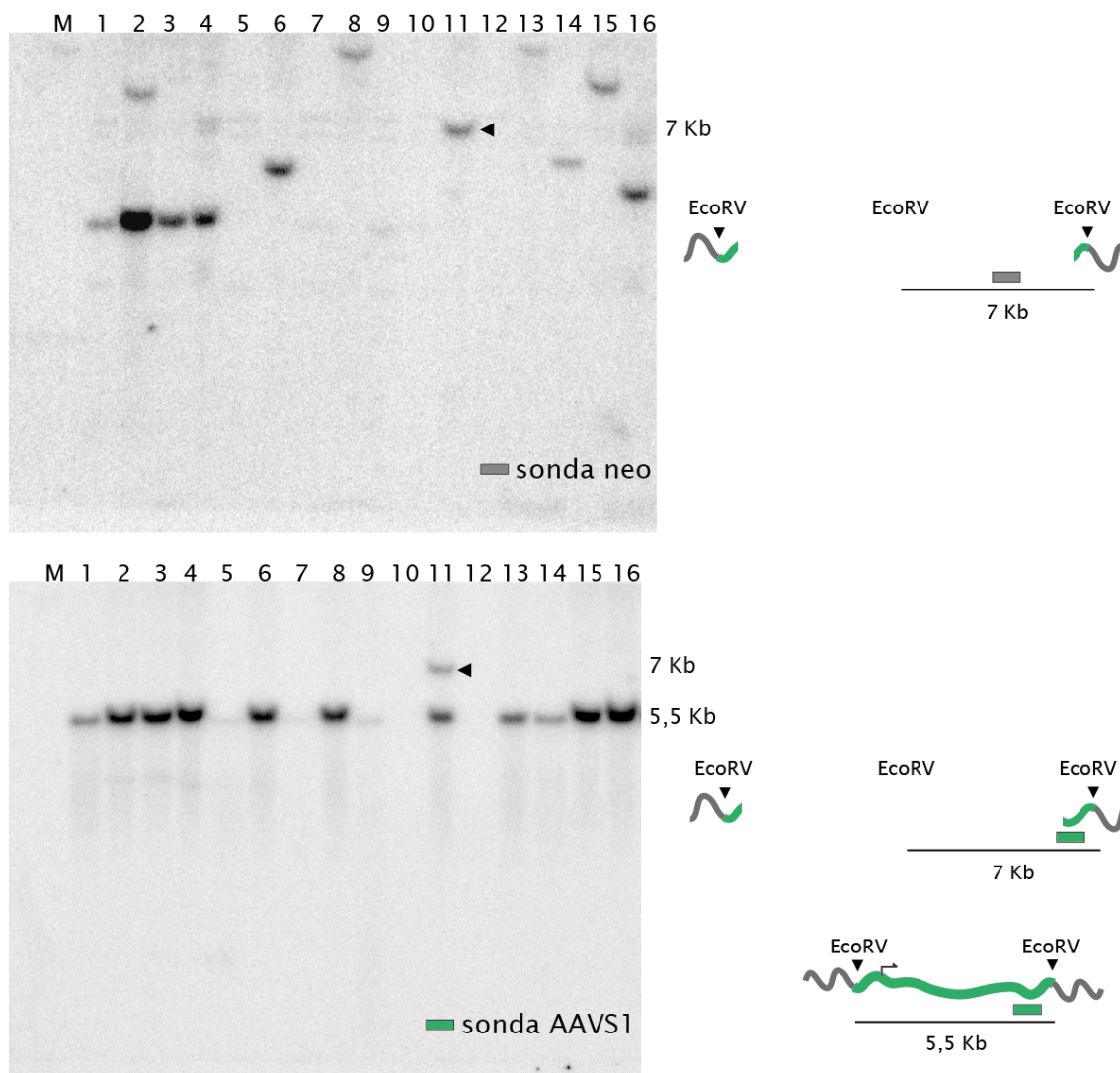
### **5.5. Caracterización genotípica. Análisis mediante Southern blot**

Debido a los resultados obtenidos, se decidió analizar clones, tanto de las placas correspondientes al ensayo 1, en el que la transducción      fue concomitante (7 clones ratio 16:1); como del experimento en que la transducción con la enzima se realizó      tras la transducción con el      reportero, ensayo 2 (7 clones ratio 16:1). Como además se obtuvieron clones correspondiente al control transducido con el sistema reportero únicamente (32 colonias), se decidió analizar células procedentes de dicho ensayo (2 clones).

Para ver la especificidad de la inserción se recurrió a análisis mediante *Southern blot*, para ello, se diseñó tanto (i) una sonda externa (propia del locus, sonda AAVS1), que pudiera mostrar las reacciones de inserción específicas en el locus AAVS1, debido a un cambio en la migración de la banda endógena; como (ii) una sonda interna de la casete (sonda neo), la cual mostrará si hay una única inserción y si es específica en AAVS1 o si por el contrario es aleatoria o múltiple. De los 16 clones (14 experimentales y 2 controles) que se caracterizaron 12 (clones 1, 2, 3, 4, 6, 8, 9, 11, 13, 14, 15 y 16) fueron capaces de proliferar con normalidad; mientras que 4 (clones 5, 7, 10 y 12), entraron en senescencia y no sobrevivieron al proceso de expansión clonal. De los 12 clones (10 experimentales y 2 controles) de los que se disponía DNA, uno de ellos -clon 11- fue positivo, por la presencia de una única banda de tamaño 7Kb, coincidente tanto con la sonda AAVS1, como con neo y no mostrando banda alguna más (con la sonda neo), que hiciera suponer inserción

aleatoria (figura 45). Cabe destacar que el clon 2 es el único que parece haber integrado varias copias de la casete , por la presencia de dos bandas y una mayor intensidad de estas.

Con todo ello, se puede estimar que se dispone de un sistema en el que hay una especificidad de entre el 14,3 y el 20% (14,3%: 1 clon positivo de 7 amplificados; 20%: 1 clon positivo de 5 viables), utilizando el sistema seriado



**Figura 45. Análisis genotípico de la integración**

. Análisis por *Southern blot* de los clones neoR aislados

. El DNA se digirió con la enzima *EcoRV* cuyas dianas se representan esquemáticamente en los dibujos de la derecha, así como el tamaño del fragmento de restricción que hibrida con la sonda indicada: (arriba) sonda neo; (abajo) sonda del locus AAVS1. La estructura del constructo tras la inserción específica en AAVS1 se representa en el formato esperado de acuerdo a la estrategia anticipada (ver figura 41). Abajo se representa el tamaño del fragmento *EcoRV* que hibrida con la sonda en caso del locus AAVS1 sin inserción. Los números corresponden a los de los clones estudiados. M. Marcador de peso molecular.



### MODELO 3

Sistema para entrega de formas extracromosómicas con mantenimiento y partición.

#### 1. Objetivo del diseño

Antes de comenzar esta tercera parte se cree conveniente recordar que el interés del presente trabajo es el desarrollo de sistemas de transferencia génica, que empleando vectores virales no replicativos, permitan la expresión génica y/o su silenciamiento *in vitro* e *in vivo* con mínima interferencia en el genoma celular.

La permanencia de un transgén en un estado extracromosómico es uno de los requerimientos imprescindibles que debería cumplir un vector ideal[30], [170], sin duda se trata de la estrategia de menor interferencia con el genoma residente. En dicha formulación génica coinciden algunos de los requerimientos ya postulados. Los sistemas episomales descritos en la literatura (introducción, apartado “modificación no-integrativa no-limitada: episomas replicativos”) basan su eficacia en responder a dos preguntas:

1. ¿qué hace que un plásmido con un origen de replicación no se integre? y
2. ¿qué elementos debe contener un plásmido para replicarse?

Su desarrollo se ha beneficiado de los avances en el conocimiento en las últimas dos décadas sobre el proceso replicativo del DNA. De dichos estudios se empieza a caracterizar que elementos génicos están implicados y cuales permiten avanzar en el diseño e implementación de episomas eficaces.

El objetivo es desarrollar nuevos elementos episomales, derivados de virus, que en ausencia de la expresión de proteínas virales para la replicación plasmídica, permitan la partición en células hijas.

2) que con una frecuencia variable, pero baja (estimada entre  $10^{-4}$  y  $10^{-5}$  [198]) pueden integrarse al azar. Dichos intermediarios de la retrotranscripción por tanto parecen servir como modelo de estudio para tratar de responder a las preguntas planteadas anteriormente.

Con este escenario planteado, se quiere explorar la utilización de episomas derivados de vectores lentivirales, modificándolos en una doble dirección:

- a. minimizando el potencial de integración inherente a todo DNA exógeno
- b. introduciendo elementos que permitan su mantenimiento (partición y replicación) sin expresión de proteínas virales, como las replicasas (LTAg, E1/E2, EBNA, etc)

Para ello se diseñaron vectores con los componentes que se comentan a continuación.

## 2. Elementos del diseño.

### 2.1. Lentivirus

Como se ha comentado anteriormente, se hará uso de sistemas lentivirales por tres razones principales:

- i. se trata de sistemas de transferencia génica con una alta capacidad como portadores de secuencias exógenas o transgenes
- ii. el proceso de la retro-transcripción genera moldes circulares que son transportados al núcleo eficazmente [123]
- iii. las secuencias contenidas en dichos episomas carecen de secuencias procariotas (origen de replicación bacteriano y genes de resistencia) y por tanto pueden considerarse eucariotas; un requerimiento descrito como determinante para la eficacia del mantenimiento episomal [208].

Tomando estos sustratos se pretende evaluar cuales de los elementos implicados en el mantenimiento episomal y que combinaciones, aportan a este modelo su potencial como episomas replicativos.

### 2.2. Orígenes de replicación

Se seleccionó para su estudio elementos o secuencias eucariotas previamente caracterizadas como *Oris*. Para ello, se ha considerado que han de cumplir una serie de requisitos, como son:

- i. secuencias de probada funcionalidad en mamífero
- ii. que la tasa de replicación que promuevan sea de una vez por ciclo celular y
- iii. que tengan un tamaño fácilmente acomodable en un esqueleto lentiviral

### 2.3. S/MAR

La incorporación en plásmidos con potencial replicativo han establecido que son los encargados del mantenimiento [212], la estabilidad mitótica [149] y un control transcripcional activo [213], [214] evitando o disminuyendo la posibilidad de metilación [215]. Por todo ello, se estimaron como uno de los componentes necesarios en el diseño del sistema episomal.

#### 2.4. Descripción del diseño

En las células transducidas se utilizará la expresión de ambos marcadores (neo y eGFP) para establecer la cinética de replicación extra-cromosómica y la partición episomal. Posteriormente, se caracterizará la estabilidad fenotípica en relación con la presencia episomal/integrada de los episomas lentivirales.

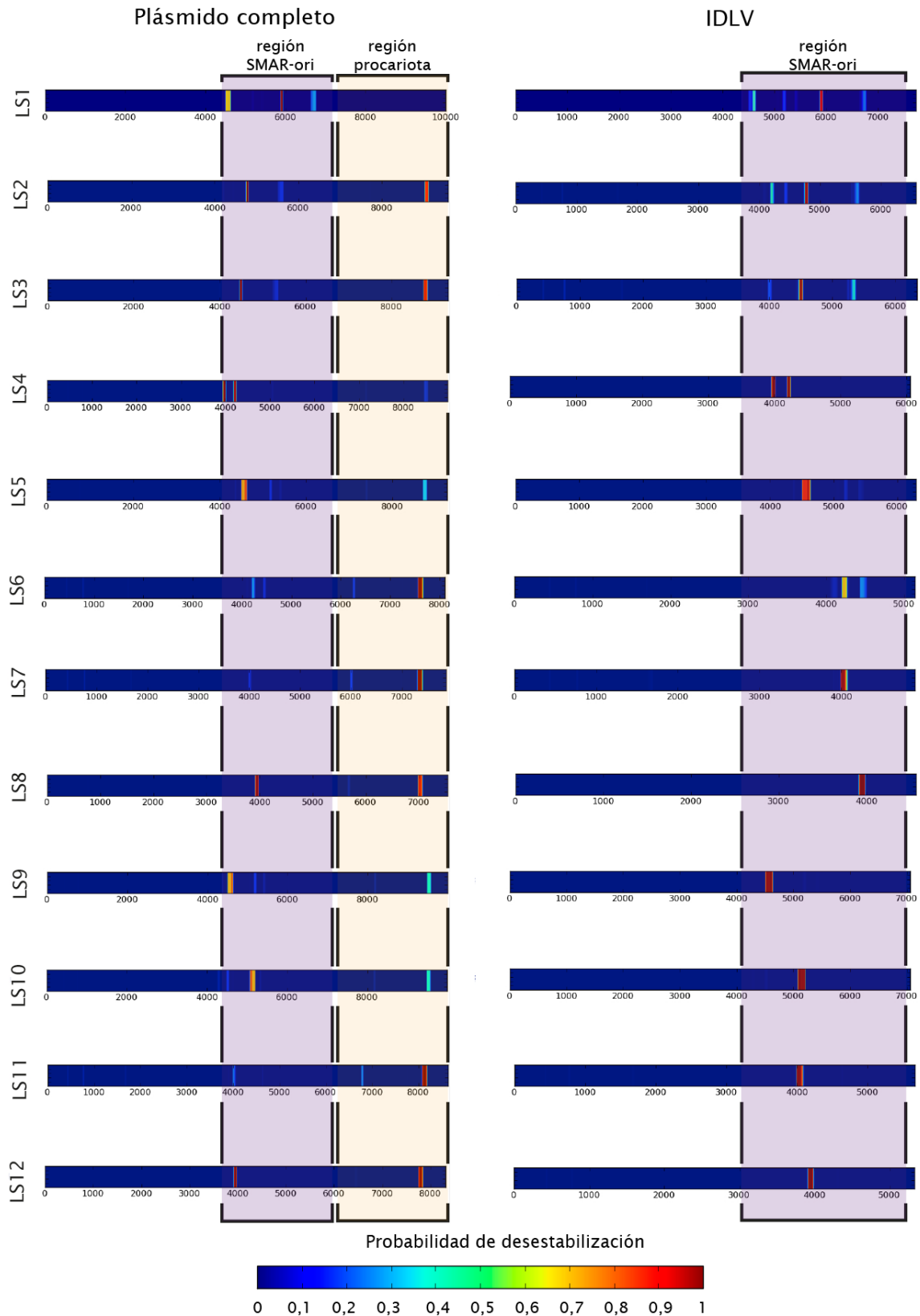
La combinación de los tres *Ori* indicados , junto con los cuatro S/MAR conduce al estudio mediante 12 IDLVs ( , figura 46B). Como controles se generaron un IDLV sin *Ori* ni S/MAR a modo de control negativo; o uno con las secuencias correspondientes al *oriP* y EBNA del virus Epstein-Bar como control positivo.

### 3. Estabilidad y viabilidad episomal

#### 3.1. Análisis estructural

En primer lugar se procedió a un análisis *in silico* de la desestabilización de la doble hélice de DNA inducida por stress (SIDD, de sus siglas en inglés). Estos datos ofrecen información predictiva de la estabilidad relativa a lo largo de una secuencia (medida como  $\Delta G^\circ$ ), permitiendo interpretar la probabilidad de encontrar regiones en configuración abierta (desestabilizada) que indican la presencia de posibles sitios de iniciación de la replicación.

En la figura 47 se muestran los resultados obtenidos al comparar en cada caso la SIDD entre los plásmidos lanzadera (conteniendo secuencias procariotas y eucariotas) y los generados por retro-transcripción. En todos ellos, se observó sobre perfiles parecidos la aparición de una zona de mayor probabilidad (tonos rojos) de desestabilización representada con menor energía libre SIDD para los episomas lentivirales; en los plásmidos completos, en la gran mayoría de ellos (2, 3, 6, 7, 8, 11 y 12) esta zona coincide con la región procariota (cuadro sombreado figura 47), por lo que la probabilidad de que el origen de replicación sea funcional es menor. Como ejemplo destaca al comparar la SIDD entre plásmido vs. LS para las construcciones 5 (LS5), 7 (LS7), 9 (LS9), 10 (LS10) y 11 (LS11); donde la probabilidad de regiones abiertas, asociada a una menor energía ( $\Delta G^\circ$ ) necesaria para la desestabilización, se ve incrementada preferentemente en las regiones asociadas a las secuencias SMAR y *ori*.



**Figura 47. Estudio de SIDD de los plásmidos** . Los rectángulos simbolizan la secuencia lineal de cada plásmido. Los numerales bajo los mismos indican la posición de los nucleótidos lineales. El código de color expresa la probabilidad de que una determinada región esté presente en estructura abierta, siendo rojo la mayor desestabilización y azul la menor. (A), se representan el estudio sobre el plásmido completo (incluyendo las secuencias procariotas); (B) estudio sobre las regiones

Este estudio permitió verificar que las regiones *Ori* incluidas en las construcciones son las de mayor probabilidad SIDD, como se había previsto; por lo que se desestimó llevar a cabo el estudio funcional como plásmidos de DNA previo al formato y centrar el estudio en los generados partir de aquellos.

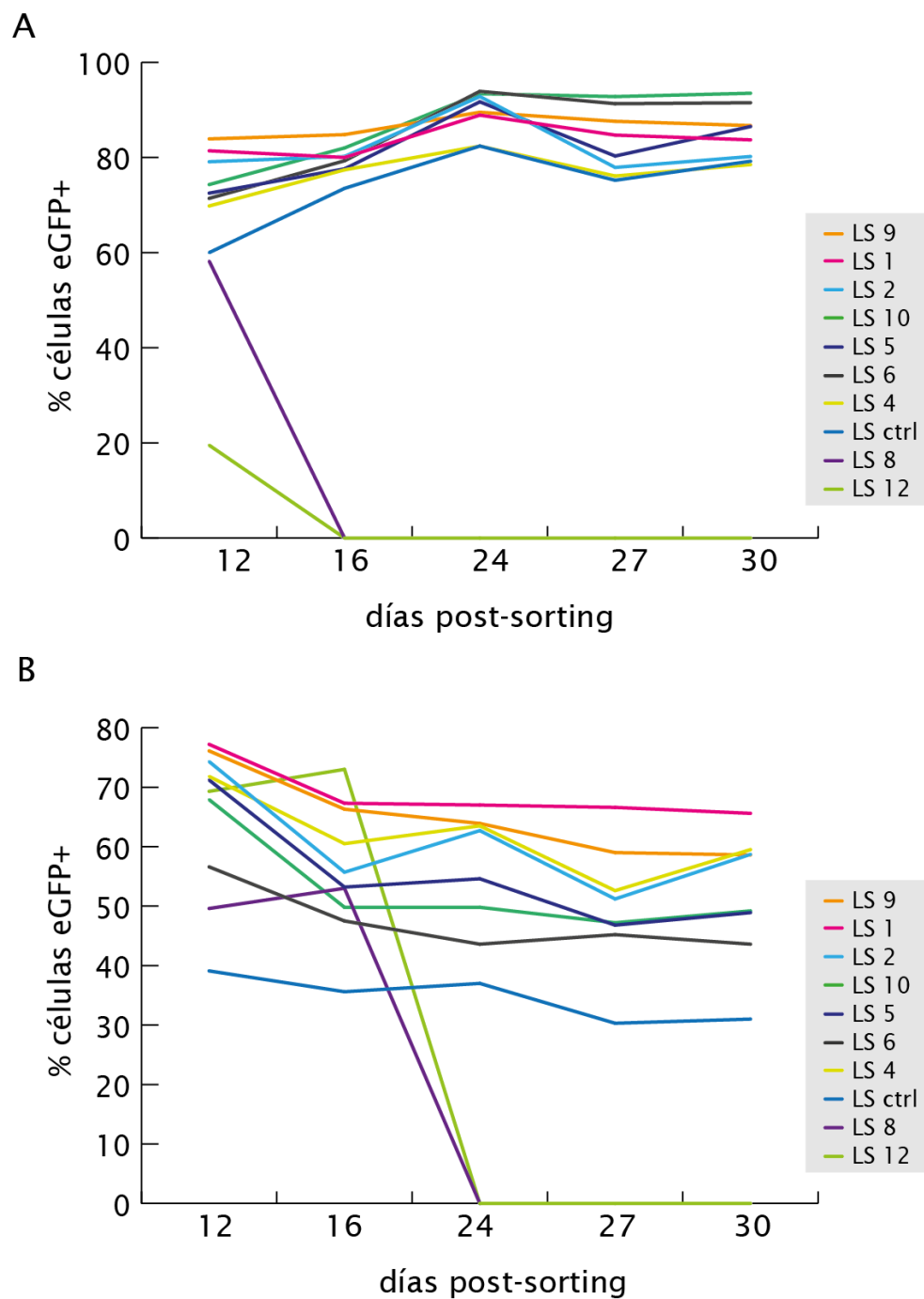
### 3.2. Producción, transducción y cinética de la estabilidad

Los lentivirus generados se emplearon para transducir células HEK293A a una MOI de 2 TU/cel. Los cultivos se mantuvieron durante cinco días para permitir el denominado establecimiento episomal (EE) [208]; es decir, el período en que el número y/o la no-funcionalidad de los *Ori*/SMAR lleva a la eliminación de células en las que ni replicación, ni partición sea operativa. Transcurrido dicho tiempo se aislaron las células que expresaban eGFP mediante separación por citometría de flujo (*sorter*) y se mantuvieron sin presión selectiva durante una semana para amplificarlos tras el aislamiento. Los cultivos se dividieron y dispusieron bien en medio de selección (DMEM-G) o en medio normal. Durante este período, una vez a la semana, se llevó a cabo un seguimiento del fenotipo eGFP+ mediante análisis por citometría. Este proceso permitirá caracterizar, a distintos tiempos post-transducción, la capacidad de replicación y segregación (partición) de los entre los cultivos con y sin presión. La figura 48 muestra los resultados obtenidos en un período de 30 días tanto para los cultivos con G418 (figura 48A), como para los que no tenían presión selectiva (figura 48B).

Lo primero a destacar es que los LS que se derivan de las construcciones 3, 7 y 11, no produjeron célula eGFP+ alguna. Incluso tras sucesivas producciones lentivirales e intentos de transducción; por lo que dichas construcciones fueron eliminadas del estudio (lo mismo sucedió con el control positivo), dada la inviabilidad de las mismas. Cabe destacar que todas ellas comparten como elemento común Por otro lado, como se podía anticipar, los cultivos en presencia del medio de selección incrementan el porcentaje de células GFP+ como fenotipo concomitante a la resistencia al antibiótico. A diferencia de ellos, los cultivos crecidos en ausencia de G418 disminuyeron levemente dicho porcentaje durante los primeros días tras el establecimiento episomal.

Tanto el LS8, como el LS12, se perdieron en los primeros pases, tanto en el cultivo con selección (figura 48A), como en el que carecía de ella (figura 48B). Los cultivos transducidos con los 1, 2, 4 y 9 mantuvieron (a lo largo del periodo de estudio) un perfil muy parecido, que conduce a una horquilla del 60-70% de células eGFP+ desde la segunda semana, sin variación significativa. Dicha tendencia, no se observa en los cultivos control (LS ctrl), si bien existe un porcentaje de células eGFP+. Por otro lado, aunque con una tasa menor de mantenimiento, se observan los cultivos derivados de la transducción con los LS 6, 5 y 10; los cuales oscilan entorno al 50%.

Para confirmar estos prometedores resultados se hace indispensable su análisis genotípico, ya que no es descartable la presencia de eventos de integración que den cuenta de la estabilidad fenotípica.



**Figura 48. Análisis fenotípico en cultivos transducidos**

fluorescencia procedente de células transducidas

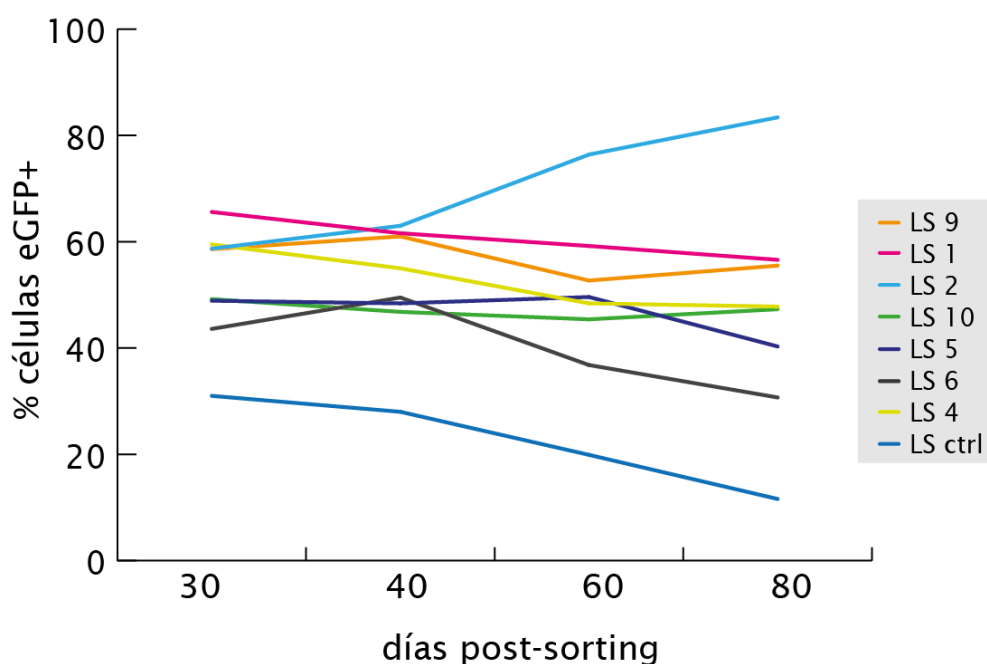
**(I).** Se muestra el análisis de la

La valoración se

realiza a lo largo de 30 días tras el establecimiento episomal, como medida del porcentaje de células eGFP+ (FACS) en los cultivos transducidos. **(A)** Representa los resultados obtenidos manteniendo los cultivos bajo selección con G418. En **(B)** no hay ninguna selección durante todo el proceso.

Antes de proceder al análisis genotípico, se decidió extender en el tiempo el estudio del % de mantenimiento episomal, pero únicamente en los cultivos sin presión selectiva (figura

49); pues parece obvio que con selección el mantenimiento es elevado (figura 48A) y probablemente se derive de la integración como consecuencia de la presión selectiva.



**Figura 49. Análisis fenotípico en cultivos transducidos** (II). Análisis en el tiempo análogo al representado en la figura 48, prolongado en condiciones de no selección, durante cerca de tres meses (80 días) desde el establecimiento episomal (7dpt).

Estos datos parecen corroborar los anteriores (figura 48), en cuanto al mantenimiento del fenotipo. Demostrando únicamente ligeras disminuciones en el % de células eGFP+; hecho anticipado, dado el carácter episomal. El aumento en el caso del LS2, requiere de un análisis o interpretación más detallada, pues parece indicar una posible integración que beneficie la proliferación del cultivo celular modificado, pero decidió no analizarse en el transcurso de la presente tesis doctoral.

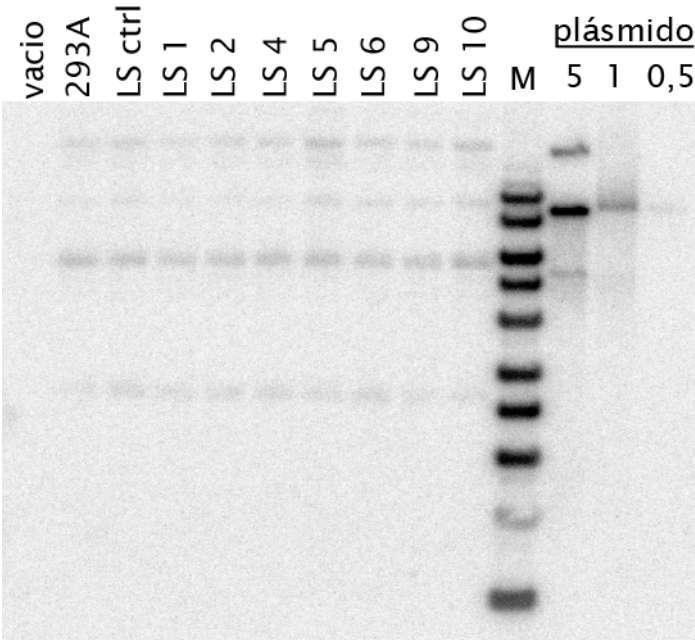
#### 4. Caracterización genotípica

Los cultivos mantenidos durante 80 días se emplearon para aislar DNA de los mismos y caracterizar la presencia de los LS transferidos. Los criterios que se emplean para establecer el estado episomal son motivo de considerable y acalorado debate[208]: *Southern blot*, amplificación por PCR, análisis de Hirt, rescate de plásmidos y análisis cromosómico por FISH (ver Discusión apartado “episomas replicativos”).

Inicialmente se realizó el estudio por *Southern blot* (Sb) sobre extractos de DNA genómico (G) y de Hirt (H, DNA de bajo peso molecular); para analizar la presencia de las secuencias LS bien en la fracción integrada (genómica) o episomal (Hirt). Sobre ambos extractos se realizó la digestión con la enzima *San DI*, que tiene una única diana en el constructo, lo que



permite el aislamiento bien de fragmentos de tamaño variable en función del sitio de integración (*SbG*) o un fragmento de tamaño definido (*SbH*)<sup>2</sup> (figura 50).



**Figura 50. Análisis de la presencia en las células transducidas** (*SbH*). Se muestra un *Sb* de muestras DNA de bajo peso molecular extraído por el método de Hirt (H), procedente de los cultivos al final del estudio (ver figura 49). La sonda empleada hibrida con el gen neo. Como control de hibridación y de sensibilidad del ensayo se incluyeron copias del plásmido crudo (carriles del extremo derecho) en las tres cantidades indicadas (ng de DNA).

Los extractos de Hirt se obtuvieron añadiendo una cantidad constante de un plásmido pCMV.ChFP al lisado celular. Esto permitió valorar la recuperación relativa en cada extracción. Para ello se cuantificó, por qPCR, la cantidad de plásmido presente en cada extracto; lo que se empleó para normalizar la variabilidad entre muestras (el valor de recuperación permitió realizar los experimentos subsiguientes con idénticos equivalentes celulares). En ninguna de las muestras estudiadas se detectó señal para una sonda específica de neo (figura 50). El análisis de la sensibilidad obtenida (carriles 5, 1, y 0,5 en figura 50) indica, que en las condiciones del ensayo, el límite de sensibilidad es de 5 - 0.5 ng/carril del gel en el *SbH*. Con estos datos se estudió numéricamente el ensayo para valorar su información.

De acuerdo a lo mencionado:

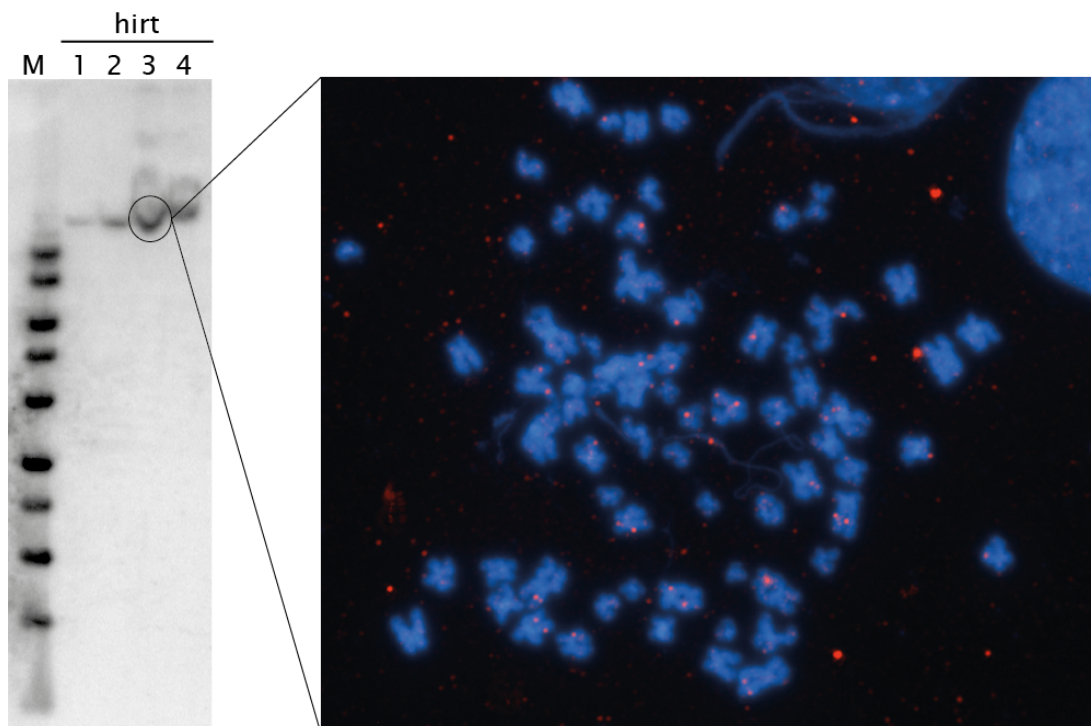
<sup>2</sup> El ensayo representado en esta figura tiene como único objeto mostrar los datos de sensibilidad que permite el estudio interpretativo posterior; ya que por tratarse de un resultado negativo no ofrece dato concluyente sobre el objeto del experimento: caracterizar la presencia integrado o no de los episomas.

|                       |  |
|-----------------------|--|
| Detección Sb plásmido | 5 → 0,5 ng   |
| Episomas (10Kb)       | 1 ng → $9 \times 10^7$ copias                            |
| Límite detección      | $4,5 \times 10^8$ - $4,5 \times 10^7$ episomas/Sb carril |

La integración de estos datos con los derivados del SbH nos permite estimar lo siguiente partiendo de que en los SbH se cargó el equivalente de DNA extraído a partir de un total de  $2 \times 10^5$  células. Dado que los cultivos presentan como promedio 50% de células eGFP+, el número de equivalente celulares con episomas cargado en el SbH sería en torno a  $1 \times 10^5$ . Bajo este supuesto, si hubiera 100 copias de episomas por células, tendríamos  $1 \times 10^7$  copias que equivaldría a 0,1ng (ver mas arriba). Si el promedio fuera de 10 o de 1 copia episoma/células, representaría entre 0,1ng y 1 pg respectivamente. En cualquiera de los casos planteados, estaríamos por debajo del índice de detección obtenido para este ensayo en nuestro sistema. Por tanto, y de acuerdo a la estimación indicada, sólo se detectaría la señal en SbH si se dispusiera de entre 10 o 100 copias de LentiSoma por célula, y se cargara en el SbH el equivalente de  $1 \times 10^6$  células (100 copias/cel);  $1 \times 10^7$  (10 copias/cel) o  $1 \times 10^8$  células (1 copia/cel).

Con el fin de verificar esta estimación, se realizó un SbH sobre células en fase temprana tras la EE (7dpt). En la figura 51 se puede apreciar, como se obtuvo una señal del extracto comparable a lo obtenido en la valoración previa (figura 50). Además, se corroboró mediante el análisis de FISH en dichos cultivos, donde se aprecia una gran cantidad de señales positivas (mayor a 50) en el núcleo de las células transducidas (figura 52). Por tanto, estos datos son consistentes, con el hecho de que los cultivos mantenidos durante 80 días post-EE deben tener entre 1 y 10 por célula.

En consecuencia el estudio de integración basado en la técnica de *Southern blot* sobre extractos de Hirt no parece suficientemente sensible y por ello se considera inapropiado, para el estudio de episomas mantenidos durante largos períodos de tiempo.



**Figura 51. Estudio de integración de episomas en tiempo temprano tras EE.** (A) *Southern blot* de extractos de Hirt (7dpt) procedentes de cultivos transducidos con LS 1, 2, 5 y 10. Se empleó una sonda LTR lentiviral. (B) El análisis de FISH del cultivo indicado (LS5), revela la presencia de múltiples copias de DNA extracromosómico específico (sonda específica de dicho LS).

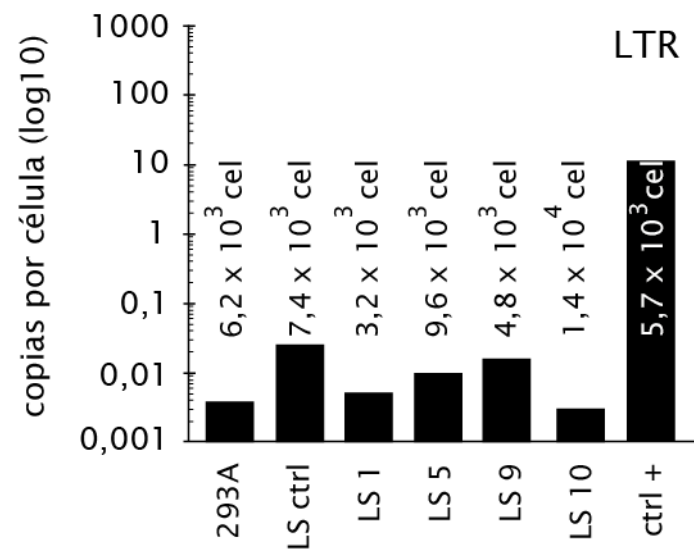
#### 4.1. Caracterización genotípica por qPCR

Al igual que anteriormente (resultados modelo 1, apartado 5.2 “análisis mediante qPCR”), se empleó la PCR cuantitativa, como metodología para detectar la presencia y número de copias de un determinado transgén; ello gracias al empleo de un gen endógeno de copia única (albúmina, accession number GI:178343). Asimismo, se disponía de un método sensible para detectar integraciones lentivirales con oligonucleótidos específicos de las LTR (ver Figura 27). Por tanto, se decidió analizar los cultivos transducidos que habían mantenido elevados niveles a 80 días post-EE del fenotipo GFP+. Para de esta manera, poder detectar la presencia de LTRs (y por tanto LS) integradas al azar en los DNAs genómicos purificados.

La figura 52 representa los datos obtenidos al aplicar esta técnica al DNAg de los cultivos transducidos

Para este análisis se han seleccionado, procedentes de los cultivos mantenidos 80 días sin selección (figura 48), clones con porcentajes variables de células eGFP+ abarcando una horquilla de >60% (LS1, LS9), 50% (LS10) y < 40% (LS5). Como control negativo, se utilizaron células parentales sin transducir y como control positivo, células transducidas con un lentivirus integrativo (LV-CMV.eGFP). Como se indica, el límite de detección del ensayo está entorno a las 0,001 copias por célula, o lo que es lo mismo, poder detectar 1 célula con integraciones, en un cultivo de 1000 células negativas. El control positivo, obtenido con lentivirus integrativo, contiene 10 copias por

célula como promedio. Los resultados obtenidos para los LS, arrojan valores menores o iguales a 1 copia /100 células (LS ctrl, LS5, LS9) o menores a 5 copias cada 1000 células (LS1, LS10), que es el valor experimental obtenido para las células parentales (293A). Se puede concluir, que existe cierta probabilidad de obtener copias integradas de los episomas (1:100 o 5:1000), en alguno de los cultivos transducidos y mantenidos sin selección, aunque dichos datos están en el límite de detección del ensayo, conforme a lo obtenido en las células control. Por ello, se hace necesario explorar en mas detalle este punto crucial en la caracterización de los LS, para dilucidar si se trata de integraciones o no.



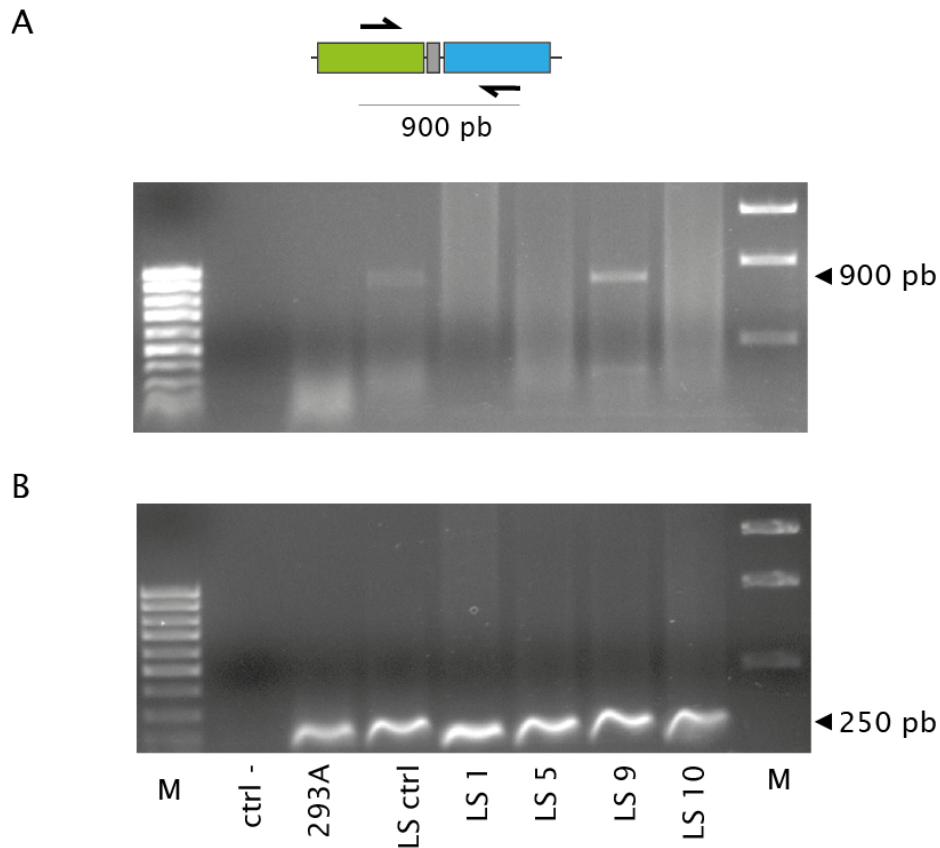
**Figura 52. Caracterización genotípica de cultivos transducidos** (qPCR). Análisis de la integración de las secuencias contenidas en los cultivos transducidos con los LS indicados. Los datos corresponden a los cultivos mantenidos 80 días en cultivo y sin selección. Se indica en cada caso, el número de células empleado obtenido mediante la cuantificación de copias del gen de la albúmina.

## 4.2. Análisis mediante PCR convencional

Los criterios de valoración de la integración han de ser cuantitativa y cualitativa. Por ello se analizó, mediante PCR convencional en condiciones de saturación (amplificación de 40 ciclos), si en el DNA genómico de las poblaciones en estudio se mantenían copias lentivirales integradas.

Se empleó una pareja de *primers* que hibridaban en la región del transgén (eGFP<sub>fw</sub>-neorev) generando un fragmento de 900-pb y para normalizar se utilizaron *primers* que hibridan con el gen de la albúmina. La PCR correspondiente a los cultivos transducidos con los 1, 5, y 10, no produjo banda de amplificación para la secuencia GFP/*neo* (figura 53A), confirmando los datos obtenidos por qPCR. El cultivo transducido con el 9, en cambio generó una banda específica, que indicaba una fracción de células en las que había habido integración. La intensidad de dicha señal era muy inferior a la

obtenida para el correspondiente control de la albumina (figura 53B); lo cual, parece confirmar el valor obtenido por qPCR (figura 52), arrojando una tasa de integración inferior a 0,01 copias/célula. El cultivo control (LS-ctrl) que tenía un porcentaje bajo pero mantenido de células eGFP+ (7% a 80d post-EE), también amplificaba una banda de 900-pb (figura 54A), indicando la presencia de un número bajo de integraciones.



**Figura 53. Caracterización genotípica de los cultivo transducidos (PCR).** Análisis de PCR con los *primers* indicados sobre la construcción génica esperada en las célula transducidas con los LS. (A) La aparición de la banda de 900pb representa la presencia en el genoma de las célula de la construcción. (B) Gel de agarosa correspondiente a la amplificación del gen de copia única de la albumina como control experimental.

Por tanto y reuniendo los datos obtenidos de ambas técnicas (PCR convencional y qPCR) se puede concluir que únicamente dos de las líneas generadas (LS ctrl y LS9) poseían integraciones; pero que estas, se encontraban entorno a 2 copias cada 100 células lo que se demuestra por la presencia de la banda de 900pb (figura 53A) y los valores obtenidos por qPCR (figura 52).

Más importante y significativo es el haber obtenido tres líneas (LS1, LS5 y LS10), que mediante las técnicas de detección empleadas, no contenían copias integradas detectables

de las construcciones transducidas. Por tanto, representa la prueba (a falta de un estudio adicional, ver mas abajo) de que sin selección, se puede establecer y mantener un sistema episomal que no se integra y se segrega establemente, a partir de una entrega lentiviral.

### 4.3. Detección episomal mediante FISH

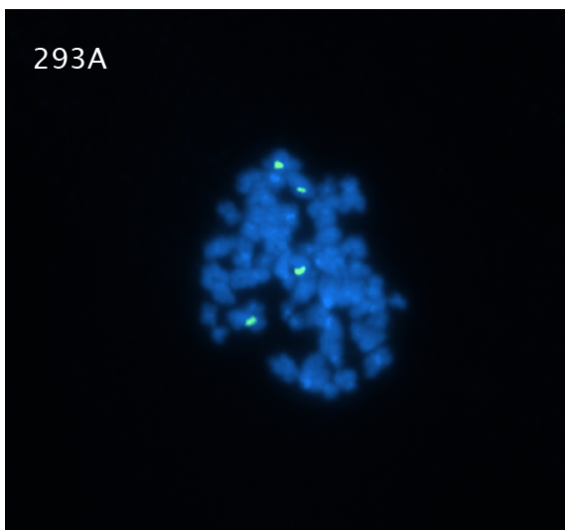
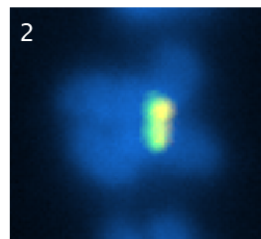
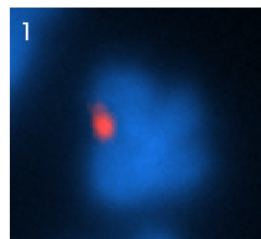
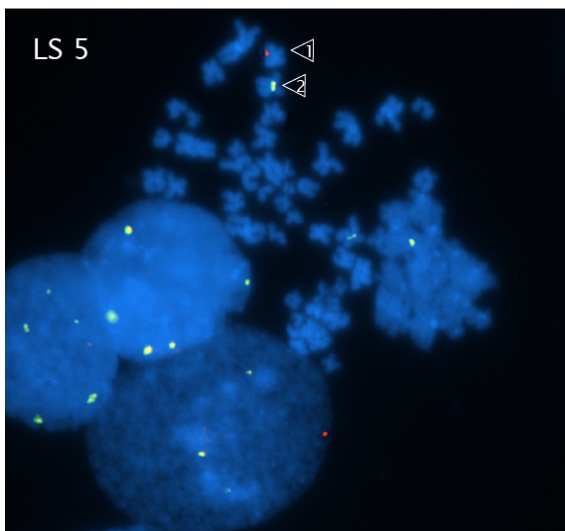
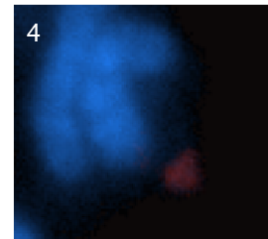
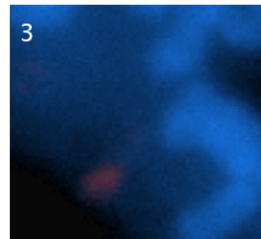
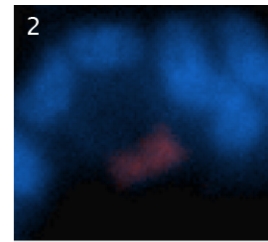
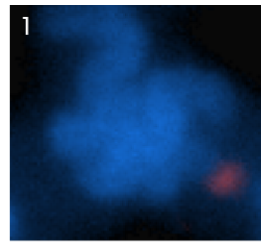
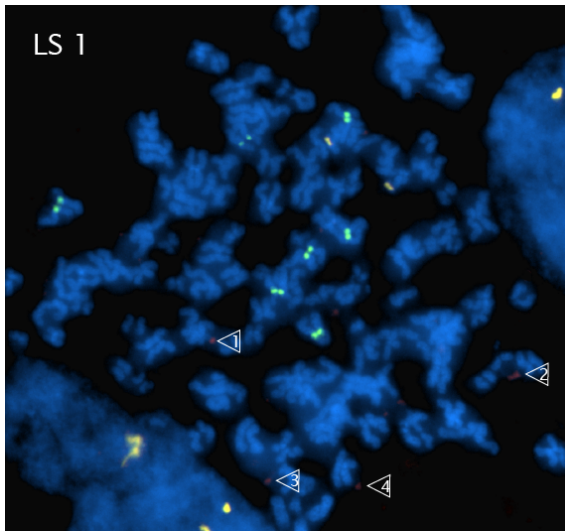
El criterio que permite confirmar inequívocamente el establecimiento del estado episomal es la visualización del DNA exógeno introducido en los núcleos. Por ello, la técnica de hibridación *in situ* con fluorescencia (FISH) se puede considerar como auténtico sistema verificador. Esta técnica permite visualizar la localización exacta del DNA introducido, pudiéndose determinar si se localiza en *spots* extracromosómicos (episomas) o por el contrario integrado en la cromatina nuclear. En la presente tesis se ha adoptado esta aproximación para confirmar los datos obtenidos por técnicas complementarias (qPCR, PCR y SbH).

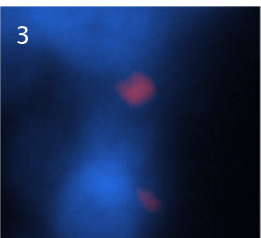
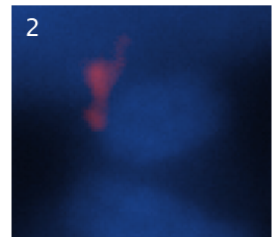
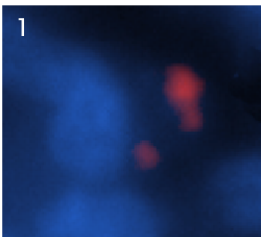
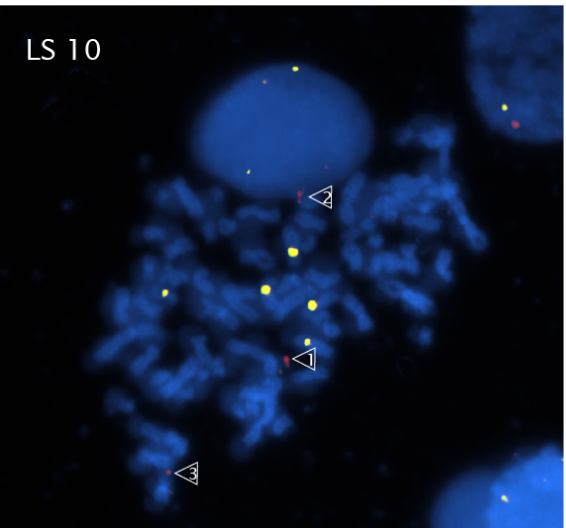
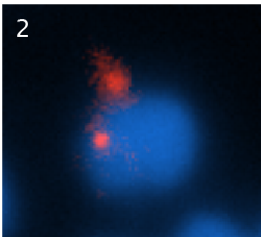
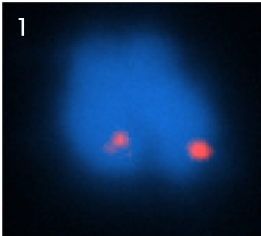
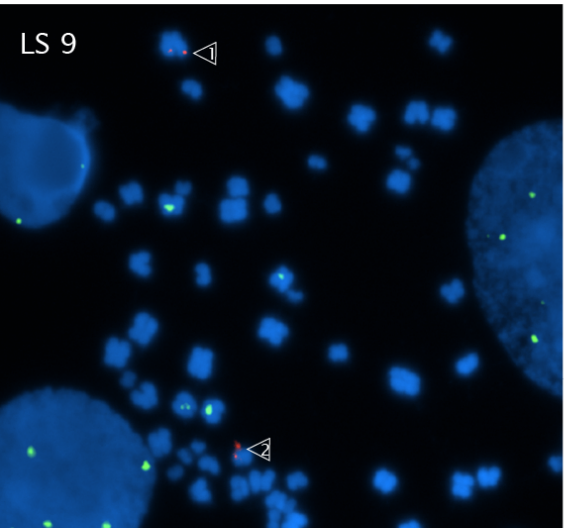
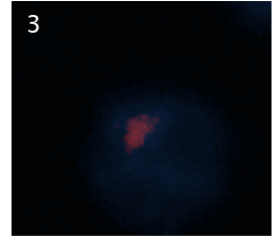
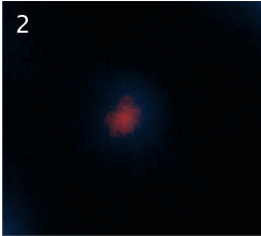
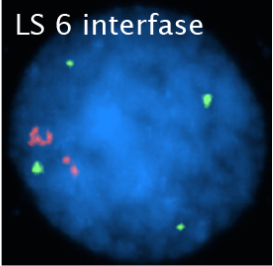
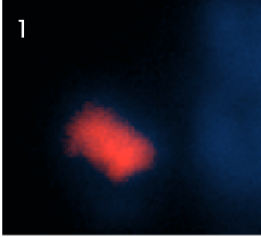
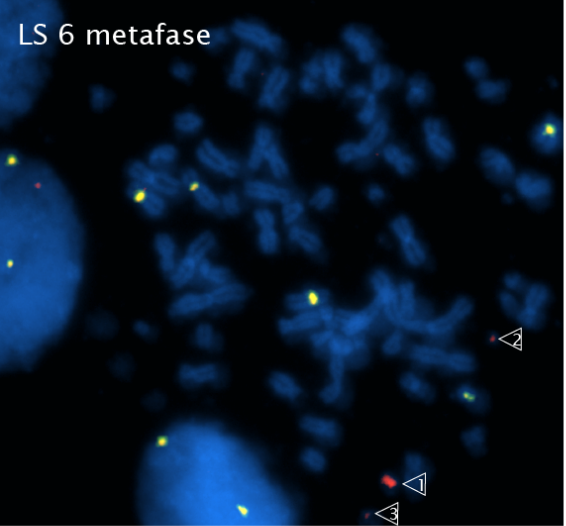
Se analizaron mediante FISH cinco de las líneas, transducidas con 1, 5, 6, 9 y 10, mediante el empleo de dos sondas específicas, una para las secuencias episomales (marcaje rojo figura 54), y otra para la región centromérica del cromosoma 3, esta última sonda se utiliza como control de detección (marcaje verde, figura 54).

En la figura 54 se muestran seis imágenes representativas de metafases hibridadas, correspondientes a cada una de las poblaciones celulares seleccionadas (5 y el control). En primer lugar cabe destacar que las células poseen en la mayoría de los casos cuatro copias del cromosoma 3 (habiendo células con tres copias, LS 5 y otras con cinco, LS 10); este hecho permitió (junto con el conteo del número total de cromosomas presentes, figura 20) interpretar la metafase LS 1, en la cual se observan 8 cromosomas marcados en el centrómero, lo que indica que son dos metafases correspondientes a dos células distintas. En casi todas las metafases analizadas, la sonda fluorescente episomal se localiza en la periferia cromosómica y/o asociado a los mismos; a excepción del LS 9, en el cual se observa una señal doble (en ambos brazos de un mismo cromosoma) que indica eventos de integración. Aunque este evento de integración no se observa en la mayoría de metafases analizadas, esta imagen de FISH se tomó como representativa dada su concordancia con los resultados de PCR (figura 53). Además la técnica de FISH permite determinar que el número de copias de las señales episomales es variable, incluso se observan células que no poseen ninguna señal. Este dato corrobora tanto el SbH (figura 50), como el análisis por citometría de flujo (figura 49).

Cabe comentar/destacar que el cultivo transducido con LS 6 presentó numerosas señales que parecen ser aglomerados episomales o podrían ser tomadas por ruido de la técnica (figura 54 LS 6 ampliación 1); es por ello que se decidió analizar núcleos interfásicos, en los que la cromatina se encuentra en un estado más relajado y lo que permitió verificar que se trataba de una señal compuesta por multitud de episomas dada su presencia en el núcleo celular (figura 54 LS 6 interfase).









**Figura 54 (páginas anteriores). Caracterización genotípica de los cultivo transducidos**

**por análisis de FISH.** Cultivos de LS 1, LS 5, LS 6, LS 9, LS 10 y control HEK293 se estudiaron empleando una sonda específica para la detección de las secuencias contenidas (rojo). Como control se empleó una sonda para marcar el centrómero del cromosoma 3. (verde). Se indica con triángulos la presencia de señales de los episomas. Los números en los mismos se corresponden con las ampliaciones mostradas a la derecha.

En conjunto las pruebas aportadas mediante qPCR (figura 53), PCR convencional (figura 54) y FISH (figura 55) son consistentes con un mantenimiento de los episomas en formato extracromosómico. Esto correlaciona el fenotipo con la ausencia de formas integradas, sugiriendo que los episomas con los S/MAR y ori ensayados (ver abajo), son capaces de replicarse y segregarse de manera estable, al menos durante tres meses en cultivo.

Estos datos ofrecen por tanto, una prueba definitiva de que los elementos presentes en los 1, 5, 6 y 10 colaboran para soportar el mantenimiento extracromosómico autoreplicativo. Los elementos contenidos en las combinaciones que hemos demostrado son:

| Nombre       |
|--------------|
| LentiSoma 1  |
| LentiSoma 5  |
| LentiSoma 6  |
| LentiSoma 10 |



## DISCUSIÓN



La transgénesis como estrategia tecnológica ha tenido un papel relevante en los avances experimentados por la biología y la biomedicina, durante los últimos treinta años. El desarrollo de la era de las “ómicas”, ha permitido que la transgénesis haya experimentado un auge como herramienta empleada en la generación del objeto de estudio, células y o animales de experimentación; con aproximaciones, bien en investigación basadas en proteómica, genómica funcional, microscopía, celómica, *high content analysis*, o en aplicaciones terapéuticas bien en estudio de fármacos o bien en la futura medicina genética basada en la terapia génica.

El empleo de sistemas tradicionales de transferencia genética, mayoritariamente integrativa, no son lo suficientemente fiables: los efectos posicionales derivados de la integración al azar y la carga génica no siempre controlable ni controlada, son dos circunstancias experimentales no deseadas aún no resueltas. Sus consecuencias, derivadas de la regulación epigenética y la interacción génica entre copias ectópicas y los loci residentes, puede llevar a procesos de reordenamientos, silenciamiento o amplificación génicas, desembocando incluso en un estadio de inestabilidad celular[220]. En este escenario muchos resultados observados a través de la transgénesis se derivan de los efectos posicionales que comprometen la expresión, por lo que no siempre es inequívocamente factible atribuir los resultados obtenidos al objeto génico de estudio y no a los mencionados efectos colaterales de la transgénesis [19]. En el extremo de lo que mencionamos, recientemente se ha demostrado que es posible generar células iPS empleando un vector control carente de los genes necesarios para promover la reprogramación[22], [221].

Hay un consenso establecido en el campo [196], el cual muestra que es necesario desarrollar e implementar, tanto estrategias como métodos de transgénesis, en los que se maximicen los controles y se desarrollen líneas isotópicas. Todo ello a fin de permitir la unicidad de los resultados y de las conclusiones; muy en particular en el contexto de las iPSC y hESC, dada su potencial en el ámbito de la aplicación biomédica.

El desarrollo de esta tesis se ha planteado alrededor del diseño y estudio de diversos métodos, basados en la combinación de tecnología existente de transferencia génica, con el fin de aportar soluciones tecnológicas en el contexto de los problemas anteriormente mencionados. Los resultados que hemos obtenido constituyen una plataforma que posibilita dichos objetivos, mediante la combinación de estrategias centradas en el empleo de IDLVs; los cuales, toman ventaja de la versatilidad en su manejo, empleo y modificación [106]. Creemos haber demostrado, en este sentido, que la permisividad de los lentivirus, los convierte en herramientas *ad hoc*; para promover la integración específica de sitio, si se combinan con los las enzimas necesarias y los elementos génicos adecuados. Dando como resultado elementos de interacción con el DNA, que son modulables en su actividad configurando sistemas bioseguros.

## 1. RMCE

La estrategia de RMCE comprende la inserción de un DNA exógeno en una localización genómica marcada previamente, mediado por la acción de una recombinasa y el reconocimiento de secuencias específicas para catalizar su acción. La utilización de recombinasas específicas de sitio se plantea como una posible solución a los problemas posicionales mencionados, pues mediante el empleo de sitios de reconocimiento heteroespecíficos permite una reacción de intercambio de casetes en sitios previamente marcados del genoma celular [222], [223]. Solventando de esta manera los problemas asociados a regulaciones tanto epigenéticas, como dependientes del entorno de la inserción. Además, ofrece las ventajas, primero de integrar una única copia; segundo, la posibilidad de producir sucesivos reemplazamientos (ya que los sitios de reconocimiento se mantienen inalterados); y finalmente, evitar la inserción de secuencias procariotas [90], [224], que se han demostrado poco o nada deseables.

Aún así, esta tecnología, como todas, no es perfecta y por tanto precisa de mejoras. Estas han de producirse en varias direcciones:

- (i) El sistema de entrega, pues principalmente se ha hecho uso de vectores o sistemas no virales poco eficientes o citotóxicos [225]
- (ii) La forma en que se controla o regula la enzima recombinasa, para de esta manera minimizar su toxicidad y los daños inducidos por esta [226], [227].
- (iii) Un sistema de identificación claro y conciso que minimice el tiempo de análisis y el volumen de trabajo. Pues aunque se trata de un método eficaz, en comparación con los métodos tradicionales de recombinación homóloga, los métodos de análisis actuales no ofrecen resultados óptimos para procesos de escalado y cribado [191].
- (iv) Elegir un emplazamiento genómico seguro, accesible y activo, en la medida de lo posible; para de esta manera solventar problemas de silenciamiento o de interferencia génica.

Teniendo en mente estos aspectos, se desarrolló el sistema Kas/Trina (K/T) (apartado resultados, figura 6): sobre una base celular marcada previamente (KAS) en un locus concreto (AAVS1), poder realizar una serie de intercambios de una manera inequívoca, controlada y segura; mediante el uso de un vector lentiviral específico (TRINA), que no sólo sirva como donador, sino también catalizador de la reacción.

### 1.1. Funcionalidad del sistema. Copias episomales

Inicialmente se verificó y validó si el diseño prototipo, entregado como sistema de plásmidos era capaz de llevar a cabo la recombinación: demostramos por primera vez, que la entrega de la recombinasa, en el mismo vector de intercambio, es factible y que el uso de sistemas de intercambio, mediado por loxP heteroespecíficos, es viable. La eficiencia fue elevada, obteniéndose entorno al 20% de eventos de RMCE (figura 9) cuantificables por la resistencia al antibiótico anticipada.

Resultó especialmente fundamental comprobar que el evento de *promoter trapping*, núcleo del sistema diseñado, era funcional y que se ponía de evidencia en las dos versiones que se propusieron, tanto mediante la obtención de células rescatables por activación de la fluorescencia, como a través de la resistencia a antibióticos. Un dato que inicialmente llamó nuestra atención, fue el hecho de la obtención de valores dispares para el rescate por los dos métodos, en cuanto a las frecuencias de *targeting* efectivo. Oscilando los valores entre el 3% ( $3 \times 10^{-2}$ ) en el caso de la activación de la resistencia (figura 7) y el 20% ( $20 \times 10^{-2}$ ) para la fluorescencia (figura 9); es decir, una diferencia de un orden de magnitud en cuanto a la obtención de clones positivos. Pensamos que esta diferencia es debida al nivel de expresión relativo necesario para seleccionar eventos positivos en cada caso: mientras que en el caso del análisis por fluorescencia, la técnica utilizada es *single-cell analysis* (FACS [228]) y permite la detección de un único evento positivo, en un amplio margen de expresión. En cambio el fenotipo “resistente” requiere niveles particulares de expresión y un *steady-state* de la enzima que en muchos casos incluso compromete la viabilidad celular, dificultando el análisis fino de aquellas células que han recombinado adecuadamente [229].

Por ello, sugerimos como una implementación futura, el uso de sistemas libres de antibióticos [229-232], siempre y cuando exista una alternativa viable, como la separación celular mediante FACS-*sorting* [231] o el uso de selección negativa (HSV/TK) [230]. Consiguiendo evitar los efectos deletéreos asociados al empleo de dichos agentes selectivos.

## 1.2. Copias integradas al azar

La verificación siguiente se realizó con el componente central del sistema: lentivirus para entrega de la casete de reemplazamiento (donadora) y recombinasa. El sistema de lentivirus defectivos en integración (IDLV) se generó *de novo* en el laboratorio, dada la patente que dificultaba la obtención del plásmido original [128]. Al igual que describen estos autores, el sistema IDLV funcionó conforme a lo esperado y su expresión en las células transducidas se perdía, paulatinamente, a lo largo del período de expansión del cultivo. Si bien, es destacable el hecho de que tanto su efectividad, generación, así como su transducción, se tuvo que optimizar, ya que los niveles de expresión son más bajos (atendiendo a intensidad de fluorescencia), la obtención de títulos lentivirales es menor (lo cual afecta a la tasa de transducibilidad) y su tiempo de visualización en el interior celular se restringe (dado su carácter no integrativo), especialmente cuando se utiliza una MOI media-baja (valores menores a 5 TU/cel) [198].

Verificamos en primer lugar cómo el sistema era efectivo sobre copias integradas al azar en el genoma (figura 11). Aunque en el diseño del sistema se utilizaron dos sitios loxP heteroespecíficos, que en teoría no deberían generar eventos de eliminación de la casete, puesto que no son reconocibles entre si (véase figura 6) [233], se evaluó el que esta posibilidad no se diera mediante el empleo de un IDLV que expresara únicamente la enzima cre recombinasa y que por tanto no entregaba la casete donadora Trina. De producirse, se podrían rescatar dichos eventos por la activación de la resistencia a higromicina al emplazarse gen de *hygroR downstream* del promotor PGK, aunque no por expresión de la fluorescencia, que sólo se produce como consecuencia del intercambio de la casete (figura 11). No se pudo rescatar evento alguno bajo estas condiciones, lo que nos

indica que dada la frecuencia no significativa (menor a  $1 \times 10^{-6}$ ), los eventos viables (higroR) son únicamente positivos a la RMCE, es decir dependientes de una recombinación fidedigna y no generados por posibles eventos de integración residual del IDLV {Banasik:2009wk}.

Está bien establecido en la literatura, que el vector de entrega y el que posee la expresión de la enzima (hasta la fecha siempre codificados en plásmidos independientes), han de estar en un formato circular, pues se favorece el recambio de la enzima y su accesibilidad a las dianas, sobre los sistemas lineales de transferencia [234]. La electroporación, es efectiva y adaptable a una gran variedad de células (a diferencia de la lipofección [235], [236]) y es por ello un sistema de elección para estrategias de entrega en transgénesis. Sin embargo provocan cortes de doble cadena o de hebra sencilla en el DNA y son altamente deletéreos para la célula [237-239] lo que dificulta la reacción de RMCE entre moldes lineales. El sistema diseñado, en cambio favorece la reacción por cuanto es probable se dé entre intermedios circulares 1-LTR y 2-LTR producidos durante la retrotranscripción [123] haciendo del sistema empleado una alternativa que debe aportar dichas ventajas. Queda no obstante pendiente un estudio sistemático comparando estas frecuencias en función del método de transferencia y la RMCE.

Adicionalmente, cuando de lo que se trata es de controlar el número de copias transferidas, los lentivirus Trina ofrecen ventajas sobre los sistemas de transfección mediada por precipitados de calcio, los cuales introducen un número altamente variable e incontrolado de copias (entre 5 y cientos), que se encuentran principalmente en concatémeros cabeza-cola [191]. Como se deriva del empleo de distintas multiplicidades de infección (figura 14), la utilización de este sistema viral aporta “flexibilidad” a la reacción de RMCE, pues ofrece la posibilidad de adecuar las condiciones de transducción y la cantidad de virus, en consonancia con los requerimientos del experimento. No hay que olvidar que el número de moldes intracelulares es una función directa del número de moldes retrotranscritos, y que estos son máximo dos por virión competente que entra en la célula.

Además, no representa una barrera la extrapolación a otros tipos celulares y su escalado según los requerimientos en la producción u obtención de clones; e incluso su aplicación podría darse en estrategias de *medium-to-high throughput screening*. Si bien los sistemas basados en transferencia de plásmidos pueden ofrecer la posibilidad de escalado, aunque con resultados más dispares tanto por las limitaciones de los métodos de electroporación, como por lipofección [91] y su extrapolación a otros sistemas celulares es técnicamente más complejo que modular la transducción lentiviral modificando por pseudotipado la susceptibilidad y permisividad celulares, pero sobre todo, aumentando la cantidad de virus, la MOI. El diseño no tiene ninguna limitación previsible para ser adaptado a otros sistemas virales [240]. Nuestros datos y experiencia, nos refuerzan las ventajas de la metodología empleada por cuanto aporta la versatilidad propia de los lentivirus, en cuanto a su pseudotipado, facilidad de manejo y amplio espectro de células transducibles y viables tras la modificación [198]. Por tanto, se cuenta con un sistema que puede ser modificado para su uso incluso en células refractarias al proceso de transfección, por medio de la adecuación del protocolo de transducción y la MOI empleada [241]. En conjunto, estas ventajas



sugieren que el empleo de un sistema viral de transferencia, para la expresión de la recombinasa, puede ser de gran utilidad ya que permite adaptar las condiciones de transducción, y así optimizar en cada caso particular eficiencia o eficacia de la reacción de RMCE.

Una de los elementos mas potentes del diseño es el disponer un sistema autorregulado de de la expresión de la enzima (gracias a su entrega en el mismo vector que es susceptible de recombinar), permite llegar a un consenso entre actividad y tiempo de expresión. Dicho factor se había conseguido optimizar únicamente para las reacciones de escisión [242], pero no para reacciones de RMCE [91]. Utilizando esta estrategia demostramos que es el mismo sistema el que limita la generación de la enzima necesaria para llevar a cabo el proceso. Esto, unido a la selección por eventos positivos minimiza el daño asociado a su expresión[243] e incluso, haciendo innecesario, el aporte exógeno de la enzima [244].

Por último, uno de los inconvenientes que presenta el sistema en la formulación que hemos trabajado , es la ausencia de un reportero que nos permita conocer el porcentaje real de *targeting*, ya que no se puede conocer el número de células que son impactadas por los lentivirus, o si el proceso de transducción se ha llevado a cabo correctamente. Por tanto, sería interesante mejorar el sistema para poder incorporar, alguna manera de visualizar la acción o la interacción virus-célula.

### 1.3. Copia única integrada en sitio específico

La panoplia de estudios sobre RMCE que se han descrito en la literatura, se han llevado a cabo sobre localizaciones cromosómicas sin caracterizar o elegidas por el mero hecho de poseer un nivel transcripcional más elevado que otros o introduciendo elementos (en la casete aceptora) que permitan su aislamiento [92], [245], [246]. Además, el empleo de líneas de ratón con una expresión constante de la enzima cre recombinasa, han demostrado un comportamiento heterogéneo (en lo concerniente a la escisión de la casete) según el tejido analizado y la posición que ocupa en el emplazamiento genómico [247].

El sitio en el que se producen las inserciones (marcaje) y por extensión las reacciones RMCE, parece por tanto de vital importancia; ya que no podemos pensar en la equivalencia de las regiones cromosómicas, no tendrán el mismo potencial transcripcional y por tanto no serán comparables desde un punto de vista cuantitativo. En consonancia con este argumento, parece indispensable la utilización de un emplazamiento genómico bien caracterizado, seguro y con un poder transcripcional activo, independientemente del contexto y tipo celular (embrionario, adulto, transformado). Por ello, inicialmente, los intentos los dirigimos a introducir la casete KAS en el emplazamiento genómico humano ROSA26, descrito recientemente como poseedor de características análogas al de ratón [195]. El problema residió, que ni en nuestra experiencia, ni a la postre en la de otros [248]; ha resultado un emplazamiento accesible para llevar a cabo la recombinación homóloga. Los datos ofrecidos por Irion y Keller no parecen tan universales en su aplicación, ni tan eficientes, como parece demostrar su artículo. Ello hizo que tuviéramos que replantear el contexto en el que estudiar nuestra aproximación.

Optamos por la utilización de *zinc-finger nucleases* (ZFN) y un emplazamiento genético seguro, que ha sido testado en sistemas tan finos y complejos como las células madre pluripotentes de origen embrionario (hESC), de origen inducido (hiPSC) e incluso hematopoyético (HSC) [63], [197]. Además, de su estabilidad y estado transcripcional activo en células primarias, se ha visto su funcionalidad en células o líneas establecidas [196]. Por tanto, se estima que la combinación de la transgénesis mediada por ZFN, con la combinación de las propiedades ya mencionadas por el sistema de RMCE empleado, proveerán un sistema eficiente y sencillo para el reemplazamiento de casetes y generación de líneas en el locus humano AAVS1.

El principal inconveniente que las ZFN, en su formulación comercial actual presentan, es la limitación del sistema de generación de clones editados, ya que la metodología se basa en un sistema de nucleofección de dos componentes (i) el RNA que codifica para la enzima y (ii) el casete a integrar en el genoma; que como se ha comentado anteriormente, son métodos altamente agresivos para las células y con una tasa de viabilidad variable según el tipo celular que se emplee [237-239]. Nosotros hemos observado dicho fenómeno (figura 17), que en el caso de determinados tipos celulares puede dar lugar a dificultades añadidas.

En este sentido, y aunque no relacionado con la toxicidad intrínseca del procedimiento, la obtención de las líneas marcadas con copia única en emplazamiento definido, no presento mayores dificultades en lo que respecta a la línea celular HEK293A. Bien es cierto, que es una línea rutinaria y de fácil manipulación y con escasos requerimientos en su mantenimiento. Por el contrario, lo concerniente a los fibroblastos primarios (BJ) fue distinto, pues obligó a la modificación adaptando métodos descritos en la bibliografía [200-202], para modificar parámetros relacionados con su cultivo, de manera que permitiera obtener clones; ya que los métodos utilizados, con la línea celular, no fueron válidos y su crecimiento sin estos factores “ayudadores” impedía su correcta y reproducible obtención.

La estabilidad de los clones, tanto HEK293A como BJ, es un parámetro fundamental en sistemas diana de subsiguientes procedimientos. El estudio que llevamos a cabo, mostró que se trata de un sistema estable en la expresión de los dos marcadores. Ya que el mantenimiento de la expresión, dio lugar a un ensayo de clonogenicidad inequívoco (figura 21) y una expresión sostenida (es el caso de los fibroblastos, figura 30); aunque como se ha podido constatar, no todas las células dentro de un mismo clon, incluso siendo isogénicas, poseen unos niveles de expresión idénticos. Es necesario considerar que la dinámica del crecimiento y su estado transcripcional están ligados. Dichos procesos que modulan la actividad transcripcional dan lugar a un ruido o fondo dentro de un mismo gen en líneas o poblaciones genéticamente idénticas, que se contemplan como sistema en oscilación replicación-transcripción [23], [249], [250].

Aunque los lentivirus utilizados son integrasa deficientes, no se pueden descartar integraciones al azar, que para la mutación utilizada y dependiendo de los estudios se han descrito entre  $10^{-3}$  y  $10^{-5}$  [73], [136], [198]. Pero, el empleo de una casete donadora que no posee promotor, unido a la ausencia de este para dirigir la expresión del gen selectivo, asegura que si ocurren, estos eventos sean indetectables; dado que se ejerce presión

selectiva a favor de aquellos eventos, donde la reacción ha sido fidedigna. No obstante, se hace necesario el estudio por qPCR para verificar la no integración lentiviral (por los efectos que pudiera introducir una inserción silenciosa como se ha indicado). Los ensayos ratificaron, que las altas frecuencias de RMCE en el sistema K/T son al menos un orden de magnitud superior a los valores estimados de inserción, y por ello no se rescatan clones con inserciones aleatorias (figura 27).

Por último, en la versión KAS2.0 la incorporación de un marcador fluorescente (eGFP), permite observar la pérdida de éste, concomitante con la aparición de los clones ChFP. Este factor fue especialmente útil, para el caso de los fibroblastos, donde la reacción de RMCE aconteció como se presuponía; pero la naturaleza de los cultivos y/o la manipulación a que se sometieron, impidió el aislamiento de clones positivos. Sin embargo, el cambio de emisión fluorescente en la población celular mixta permitió apreciar, cómo había células que habían perdido la marca eGFP+ y eran rojas (ChFp+), lo cual quedó revelado en su caracterización genotípica por PCR. Queda por determinar si mediante el empleo de métodos de selección menos agresivos (dilución límite) o más efectivos (separación celular), dicha población podría ser aislada y amplificada, dando de esta manera una alternativa para su aplicación en sistemas de cultivos primarios y por tanto más delicados, como ya se comentó al principio de este apartado de la discusión.

En cuanto a la expresión limitante de la enzima, factor que se ha postulado clave a la par que deletéreo. Se estimó la posibilidad de su empleo en dos IDLVs separados, llegándose a la conclusión de que si bien, podía ser optimizado el sistema, los datos obtenidos no eran claramente beneficiosos frente al sistema de entrega en único IDLV; es más, el sistema puede verse limitado por la ausencia de autoregulación de cre, derivando en la menor obtención de clones (o como se aprecia el aspecto “menos saludable” de estos) dado que la síntesis de la enzima es prolongada y por consiguiente permanece disponible intracelularmente, por lo que la reacción puede ser revertida (siempre y cuando el *steady-state* del molde o intermedio episomal lentiviral lo permita). Otro factor que hemos observado desaconsejaría la utilización en dos vectores para este sistema es el hecho de que las células han de ser co-transducidas y por tanto la carga viral a de ser mayor, en casos en los que la sensibilidad del sistema celular sea mayor que el observado con HEK293A.

El conjunto del proceso de RMCE basado en nuestra estrategia es altamente eficaz tanto en los tiempos como en la eficiencia con que se obtiene las células por RMCE. No obstante, como en todo proceso, desde la obtención de clones KAS correctamente impactados, hasta la verificación de los clones TRINA por su análisis fenotípico y genotípico, existe un cuello de botella: la correcta caracterización y obtención de clones KAS. Ello supone, dependiendo del tipo celular empleado, entre 40-60% del tiempo invertido. En cambio, los subsiguientes pasos han sido optimizados de forma que el volumen de trabajo directamente relacionado con el número de clones a aislar, y a su *screening* posterior, es mínimo, gracias a la eficacia del sistema de rescate de clones RMCE positivos según todo lo mencionado anteriormente (el porcentaje de éxito obtenido es 6 clones de 6 analizados). En definitiva, como se ha mencionado, el sistema es altamente

eficiente, aún empleando multiplicidades de infección bajas, pues sólo se rescatan clones TRINA positivos.

La plataforma diseñada, presenta las ventajas ya mencionadas de eficiencia y eficacia. No obstante aún hay mejoras futuras que podrán introducirse para su optimización, incrementando su versatilidad y polivalencia: (i) aunque el sistema está diseñado para llevar a cabo un único reemplazamiento, una de las mejoras que se debería aportar en sucesivas versiones es la implementación de un método o sistema para llevar a cabo sucesivos reemplazamientos [251] o incluso la escisión de la casete si el tiempo de expresión ha de ser limitado [252], evitando efectos innecesarios y la interferencia con el genoma [220], [253], que no pueden ser alcanzados por los sistemas lentivirales “convencionales”; y (iii) por último, el uso de distintas recombinasas, como flp, la cual se ha postulado como una gran alternativa debido a la ausencia de sitios pseudo-reconocibles en el genoma celular mamífero y con una actividad mejorada, por implementaciones en su funcionalidad y en el empleo de sitios heteroespecíficos [91], [191], [222].

## **2. Sistema de entrega de proteínas con interés biológico**

Como se anticipó en la introducción, existe una variedad de aplicaciones, en que el interés reside en la entrega de proteínas, tanto con una finalidad terapéutica [163], como científica [254], [255]. Este hecho, que se conoce como transducción proteica, puede llevarse a cabo por medio de la entrega de la proteína mediante su fusión con dominios virales [159] o celulares [160]; pero ello se deriva en toda una problemática asociada al coste, purificación y/o modificaciones necesarias para su correcto funcionamiento [161]. Además, aún solventándose estos problemas, siguen sin ser sistemas eficientes a la hora de atravesar la membrana plasmática y evadir la degradación posterior en el interior celular [158]. Sin mencionar los problemas de control y reproducibilidad.

Otra de las alternativas reside en el hecho de poder co-empaquetar<sup>1</sup> estas proteínas en cápsidas virales, para solventar el problema asociado a su entrada y degradación. Hasta la fecha, se han hecho diversos intentos, mediante fusiones a genes virales y aprovechando la capacidad de los sistemas retro/lentivirales de portar una gran variedad de proteínas en diverso número, co-empaquetadas en el virión junto al genoma viral [105]. Aún así, existe toda una problemática asociada a la utilización de la proteína para la entrega; es el caso de la proteína *vpr* (ampliamente utilizada), la cual se ha descrito que tiene carácter citotóxico y de parada de ciclo [165], [166]. Lo que puede tener efectos negativos tanto en el sistema de producción como en la célula que debe recibir la proteína.

---

<sup>1</sup> La palabra co-empaquetar no existe en castellano. La empleamos por lo expresivo que representa. El sufijo con- adopta la forma co- en ocasiones y por ello se ha elegido.

Otra de las aplicaciones en las que se probó las capacidades del sistema fue la valoración de la escisión mediada por cre sobre una línea celular previamente generada con una casete flanqueada por secuencias de reconocimiento para la recombinasa [206]. Está ampliamente caracterizado el hecho de que la expresión continuada de la recombinasa provoca efectos adversos e incluso efectos genotóxicos en las células, que pueden en una mal función enzimática (actuación en emplazamientos erróneos distintos del previsto) [180]. Como se comentó, los métodos basados en la entrega de DNA, aunque de manera transitoria [240], [252] no pueden solventar este problema por completo, ya que existe el riesgo residual de integración al azar. Las alternativas de transfección de RNA y transducción proteica, aunque transientes, son altamente variables[204]. Sin embargo, el desarrollo

se ha mostrado eficaz, consistente y transiente en cuanto a la entrega proteica, e incluso podría alcanzar tipos celulares altamente refractarios, gracias a la capacidad de pseudotipado lentiviral [258]. Se observó que la enzima mantiene su actividad como una función de los valores de MOI empleados, con lo que se llegaron a alcanzar más de un 60% de las células modificadas por la acción de la enzima entregada (figura 40). Por último, hay que mencionar que la línea celular empleada se generó por transfección transitoria de un plásmido y la posterior selección de una población celular mixta conteniendo distintas inserciones por célula; por tanto cabe la posibilidad de que los datos obtenidos estén infravalorados, ya que es conocido el efecto que tiene el empleo de precipitados de calcio de promover la formación de concatémeros [235], [238] y por tanto el empleo de la enzima podría estar actuando y dando lugar a reordenamientos, sin necesidad de cambio en el fenotipo fluorescente observado en la célula. Ello nos invita a generar y utilizar en el futuro, una línea celular con inserción única, para de esta manera ser capaces de tener un valor absoluto y más certero de la funcionalidad del sistema.

## 2.1. Aplicación al desarrollo de híbridos

Por último se empleó el sistema de entrega de proteínas para caracterizar la funcionalidad de la enzima rep78 de AAV en la integración sitio específica en este formato de entrega. Diseñamos y validamos en primer lugar un método de análisis y cuantificación de dicha actividad de manera inequívoca. El sistema novedoso que hemos denominado sistema reportero de *splicing* (SRS), se fundamenta en el correcto reordenamiento de los exones contenidos en una construcción de forma direccional respecto a un promotor también incluido en la misma (figura 41), lo que genera células eGFP+ y resistentes a neomicina. Dicho reordenamiento se dirige por la interacción de las regiones implicados en la reacción de rep78, para promover la integración en AAVS1 [207]. Los datos obtenidos indicaron un resultado altamente positivo, dado que se obtuvo un gran número de clones celulares. No obstante, aunque en una frecuencia muy baja, aparecieron colonias G418R en los cultivos control transfectados con el vector (SRS) y sin enzima. Atribuimos la aparición de estos eventos falsos positivos (no eran fluorescentes mediante microscopia), al hecho de que la integración residual [73], [136], [198], pudiera coincidir con una entrada en el genoma “*in frame*” con la expresión de un gen endógeno celular;

Por el contrario, la señal de fluorescencia sólo puede ser rescatada en el supuesto de la integración mediada por la enzima y a través de su actuación sobre el elemento , aunque la integración se produjera en otro lugar en el genoma o sitio pseudo-AAVS1 [41]; hecho que no ocurre, pues sólo cuando los plásmidos expresando rep78 y SRS se co-transfectaron se obtuvieron células fluorescentes.

En cuanto al sistema de integración dirigido, en una primera tentativa no se consiguieron obtener células positivas mediante su análisis por citometría de flujo. Este hecho pudo ser debido a que la proteína rep78 resulta inhibitoria en la producción y por tanto se hace necesario su separación , (1) una que encapside la proteína, dado que el proceso de encapsidación puede generar partículas VLP las cuales no necesitan de un genoma para su formación y (2) otras partículas que sean las portadoras del genoma a entregar. Otra posible explicación que se encontró a este experimento fallido fue (ii) la ausencia de concomitancia en el núcleo celular entre los componentes requeridos, es decir, enzima y sustrato ya que los intermedios se han descrito como óptimos a partir de las 24 horas post transducción [123], por lo que la solución propuesta de separación se hace más idónea, para poder llevar a cabo una estrategia de transducción seriada.

Con esta segunda aproximación, se obtuvieron resultados prometedores, ya que el número de clones fue muy superior al de los controles, y su análisis nos dejó una eficiencia del 15% (integración sitio específica) después de la transducción del sistema reportero SRS. Aún así, otros sistemas han dado eficiencias de entre el 17% [263] y el 48% [264], si bien utilizan métodos de transfección y electroporación que, como mencionamos con anterioridad, tienen efectos no deseados en el genoma celular.

En este contexto, podría ser mejorable la estrategia de transducción,

Una de las alternativas que se debería emplear es la selección por el doble marcaje, resistencia a antibiótico y fluorescencia, pues de esta manera sólo se rescatarían clones en los que la inserción ha ocurrido mediada por rep78 ; y no como se estimó anteriormente, por la posibilidad de una entrada en el marco de lectura del ORF del gen de resistencia a neomicina en el genoma celular. Lo que se podría denominar evento de *gene trap*. Aún así, los estudios previos han estimado, que los clones obtenidos en sitio no específico pueden ser debidos a la integración descrita en AAVS2 y AAVS3 [41]; hecho por demostrar en el sistema que hemos empleado, pues no se han utilizado AAV2wt con las ITRs intactas, como es el caso del estudio que las describe.

Hasta la fecha, las aproximaciones de utilización de distintos sistemas virales híbridos ha pasado por tratar de aprovechar el carácter específico en la integración de AAV, pero mediante el empleo de vectores más complejos y difíciles de manejar como los herpesvirus [265-268], para los que no resulta fácil la modificación de su tropismo. Además, en nuestro conocimiento y hasta la fecha, los híbridos generados representan

que hemos demostrado de elevada eficacia, y si bien es mejorable, posee características prometedoras para ensayos de terapia génica por ganancia de función.

como ocurre con otras enzimas [71], [76] o en el caso de la recombinación homóloga[269]; ampliándose así la disponibilidad para transgenes o genes terapéuticos a testar. Incluso y permitiéndonos ir más lejos, está el hecho de que se trataría de un sistema idóneo para la generación de líneas isotópicas de manera rápida, eficaz y sin el excesivo coste de las estrategias comerciales existentes (ZFN y meganucleasas).

Aún así, el sistema tiene ciertas restricciones, como el hecho de que únicamente se puede hacer entrega de material génico en el locus AAVS1, dada la especificidad de rep78[58]. Esta circunstancia abre la puerta a nuevas opciones en su aplicación como puede ser el hecho, aun por evaluar, de realizar la entrega de ZFN [71], [197] o meganucleasas [76], [77], permitiendo de esta manera una integración a la carta. O retomando nuestros datos anteriores, su utilización en sistemas RMCE, para hacer la entrega de la enzima cre y dejando una mayor capacidad de carga de material génico por parte del vector de transferencia. O como han destacado otros estudios, testar la alternativa de ser usado como un sistema terapéutico, bien sea por su posibilidad farmacológica controlada [163], como por su posible aplicación al campo en expansión de la reprogramación mediante el empleo de una tecnología limpia y eficaz [270].

### 3. Episomas replicativos

La expresión prolongada de transgenes puede realizarse a través de la integración permanente del vector de transferencia en el genoma celular, o mediante la retención en el núcleo de una unidad extracromosómica capaz de replicarse y segregarse, lo que se denomina un episoma.

La cinética y estabilidad de dicha expresión en episomas en células en división o tejidos con alta tasa de renovación/división depende de una eficiente retención del vector, lo cual puede ser logrado mediante el uso de vectores episomales replicativos. Estos vectores ofrecen numerosas ventajas sobre los sistemas integrativos, debido al mantenimiento del transgén en un estado extracromosómico [271], evitándose los efectos posicionales, y sin afectarse de manera directa a los genes celulares, [19].

Como se comentó en la Introducción, la gran mayoría de los sistemas episomales replicativos estudiados y desarrollados (a excepción de los recientemente descritos derivados de pEPI [148]), están basados en sistemas de origen viral {SJaneFlint:uh} y que necesitan la actividad de sus propias proteínas (large T antigen, EBNA, etc) para el mantenimiento y la replicación; las cuales tienen o poseen carácter inmunogénico y oncogénico [139].

El reto del desarrollo de estos sistemas está en conciliar la replicación *once per cell cycle* (OpCC), como los orígenes eucarióticos de la célula y la segregación con la cariocinesis. La gran pregunta es ¿qué hace que un DNA exógeno no se integre? ¿qué debe contener para ser replicado una única vez? Y la respuesta se haya, al menos parcialmente, al incluir ori en el DNA. El estudio del proceso replicativo celular ofrece la disponibilidad de algunos de estos elementos bien caracterizados. La segunda parte, dividirse OpCC y segregarse parece descansar en las secuencias S/MAR de unión a cromatina y retención nuclear, verdaderos organizadores genéticos de actividad transcripcional y, ahora sabemos que también de la replicación celular. Incluidas como secuencias en un plásmido se ha visto puede favorecer el mantenimiento y los niveles de la expresión gracias a evitar el silenciamiento del transgén [177].

Excelentes trabajos del grupo de G. Bode [177], [208], [212] han esclarecido algunos conceptos para favorecer la funcionalidad de los episomas. De entre todas, nos merece especial atención la importancia de carecer de secuencias procarióticas: su presencia promueve la eliminación por la célula [208]. Los minicírculos son la expresión de estas ventajas [151].

Nosotros hemos tomado como concepto la presencia de episomas eucarióticos derivados de lentivirus:

De hecho, el empleo como vectores de terapia génica se fundamenta en estos círculos, si bien sólo aplicables a sistemas de baja tasa replicativa [128]. Y hemos estudiado la manera de convertirlos en replicativos, capaces de partición, y por tanto capaces de emplearse en sistemas celulares con una alta tasa de recambio (hematopoyesis por ejemplo).



Para ello, se ha estudiado de manera sistemática la combinación en vectores de secuencias S/MAR y ori, no caracterizadas en conjunto previamente.

Los resultados obtenidos han demostrado que algunas combinaciones son los mas ventajosos y consistentes en nuestro sistema para entregar vectores episomales replicativos basados en vectores lentivirales deficientes en integración, . Por tanto, concluimos que son una alternativa factible a los métodos empleados hasta la fecha; ya que conseguimos un mantenimiento episomal en ausencia de presión selectiva durante casi tres meses. Todo ello gracias a la unificación de gran parte de los criterios necesarios para el empleo de vectores episomales, esto es:

1. una entrega eficiente en gran variedad de tipos celulares,
2. una expresión génica sostenida,
3. simplicidad en su manipulación *in vitro*; que unido a los elementos necesarios (ORIs y S/MAR) además ofrecen
4. retención nuclear,
5. no integración,
6. segregación celular y
7. tasa de mantenimiento alta.

Hasta la fecha se habían descrito sistemas que carecían, de manera constante, de alguna de estas características; por ejemplo, si nos fijamos en los sistemas virales, era necesario el empleo de proteínas con potencial oncogénico [272], [273], su manejo resulta complicado, van desde la carencia de pseudotipado y por tanto reducción en el abanico de células a emplear [274], hasta la posesión de una replicación incontrolada, que puede conllevar a reordenamiento [275]. Con respecto a los sistemas no virales, en teoría cuentan con la ventaja de poseer una capacidad ilimitada de tamaño y una menor inmunogenicidad; pero se ha demostrado que su eficiencia de entrega (de por si baja o menor que con los sistemas virales), cae al aumentar el tamaño del vector [276]. El problema de la entrega, es compartido con otra de las grandes promesas en el ámbito de la terapia génica, los cromosomas artificiales, cuya utilización se ve restringida a determinados tipos celulares y en sistemas *in vitro* [277, 278], sin contar el hecho de que los sistemas de entrega provocan la rotura del DNA y por tanto su propensión a ser potencialmente degradados o integrados por la maquinaria celular [279], lo que conllevaría complicaciones *in vitro*, pues haría necesaria la selección o la separación celular (*sorting*) y no sería aplicable su uso *in vivo*, dada la baja eficiencia; que conllevaría problemas a la hora de corregir una deficiencia genética.

La principal dificultad la ha representado la determinación de la no presencia de integración. El principal obstáculo al empleo con garantías de no interferencia. La integración ha de estudiarse bien sobre DNAg (integrado) o DNA de bajo peso molecular (Hirt's; no integrado). La no detección se convierte en un escollo experimental que debe acompañarse de controles exhaustivos, unido al hecho de que un bajo número de copias entra en conflicto con el límite de detección de las técnicas disponibles. Con todo, hemos aplicado los sistemas de análisis del mantenimiento episomal: *Southern blot*, PCR, convencional,

qPCR y análisis mediante FISH. En este sentido es fundamental entender que la presencia de S/MAR, pueden muy bien contribuir a la co-purificación de poblaciones no enteramente puras de ninguna de las dos [208]. Todo ello se ha tenido en consideración para el estudio de los datos que han validado el modelo y lo podemos resumir en:

1. El sistema de PCR, nos da una visión parcial de la ausencia o presencia de integración. Combina a la perfección pureza DNAg y H, con sensibilidad. Incluso cuando se obtiene banda, como es el caso el clon LS 9 (figura 54), podemos plantear el conflicto de la co-purificación mediada por S/MAR [208], por ejemplo.
2. En el análisis por *Southern blot*, se ha podido constatar, que los resultados obtenidos no son concluyentes, puesto que la muestra que se maneja tiene un número de copias (en el supuesto de haberlas) inferior, en al menos un orden de magnitud, al nivel de detección que nos ofrece la técnica. Es por ello que sólo realizando el análisis sobre un número mayor de equivalentes celulares se podrá tener en cuenta el valor obtenido del ensayo. Además, el protocolo de Hirt es efectivo en los primeros estadios del establecimiento episomal, cuando el número de copias es alto (figura 51); pero dicha efectividad se ha visto claramente reducida con las sucesivas rondas de replicación [212].
3. En el caso de la PCR cuantitativa, se cuentan con niveles de detección mayores a los empleados por el *Southern blot*, y utilizando una curva patrón correctamente generada y los controles pertinentes, arroja datos más fiables que las técnicas anteriores, incluso de inserciones únicas en el genoma [280], y los datos manejados a lo largo de la presente memoria.
4. Por último, la técnica de FISH, que mediante el marcaje con una sonda específica de nuestros episomas, permite dilucidar donde está (en el contexto nuclear y cromosómico) el objeto de nuestro estudio y visualizar si está integrado (dos marcas en cada una de las cromátidas de un mismo cromosoma), o si está en asociación con la cromatina (una única marca, en la proximidad cromosómica) o si está libre en el contexto nuclear (marcaje separado de los cromosomas) [149]. Con lo cual se confirma la ausencia de integración y por tanto, se dispone, no solamente de un dato numérico (qPCR), sino además de un emplazamiento físico visible.

#### 4. Futuras aproximaciones

El desarrollo de la presente tesis doctoral se planteó inicialmente como tres apartados conectados pero no interrelacionados. No obstante, según avanzamos en su estudio y desarrollo se abrió ante nosotros una nueva posibilidad: ¿por qué no relacionarlos estrechamente? Al final creemos haber sido capaces de fusionar

Realmente, las opciones que estos sistemas nos ofrecen son muy amplias, tanto a través de sus combinaciones, como por separado. Se trata de sustratos ideales para el empleo en ensayos biomédicos, tanto de generación de modelos celulares (que mimeticen enfermedades); como para corregir genéticamente células con una determinada patología (empleo de transgénesis relativa a enfermedades monogénicas), tanto en formato episomal, como integrados en localización concreta y segura. No sólo hemos de restringirnos a estas aplicaciones, sino que se abre ante nosotros una serie de herramientas con un potencial de aplicación enorme, en biotecnología (producción de enzimas, proteínas recombinantes, anticuerpos, etc.) o en investigación (permitiendo la generación de datos altamente reproducibles, de interés en análisis multicéntricos y donde los controles sólo sean lo que se espera de ellos).

Con todo ello, por qué no permitirnos el lujo de fantasear o soñar, pensando que el día de mañana estas herramientas pudieran ser tan rutinarias como hoy día lo son las enzimas de restricción, los plásmidos bacterianos o los mismos lentivirus.



## **CONCLUSIONES**



1. El diseño del sistema dual para intercambio de casetes mediado por recombinasas (RMCE) dirigida por lentivirus no integrativos (IDLV) y que hemos denominado K/T permite obtener eventos de recombinación fidedigna, basado en *promoter trapping*, con una elevada eficiencia y simplicidad.
2. El sistema K/T es eficaz entre (a) copias donadora y aceptora no integradas en el genoma, (b) copias aceptoras integradas al azar en el genoma y (c) sobre copia única aceptora integrada en sitio específico del genoma de células humanas.
3. La entrega auto-limitante (IDLV-TRINA) de la recombinasa cre es un diseño funcional que expresa transitoriamente la enzima. Modulando tanto el nivel, como la temporalidad a modo de maximizar la reacción de RMCE y minimizando su genotoxicidad intrínseca.
4. Las líneas modificadas KAS muestran una estabilidad génica al menos durante 30 ciclos de duplicación; demostrable por la permanencia de una copia integrada en el locus AAVS1 y por el mantenimiento del fenotipo transferido con dicha construcción génica.
- 5.
6. El procesamiento de los ORF heterólogos cre y rep<sup>78</sup> no interfiere en sus actividades enzimáticas. La actividad específica entregada a través de dichas partículas lentivirales, es suficiente para que se produzcan, de forma eficiente, los procesos de corte y recombinación entre fragmentos de DNA. Ello es modulable por la multiplicidad de infección.
7. La incorporación
- 8.

9. Los intermedios circulares pueden adquirir competencia en replicación y segregación nucleares, si se introducen ciertas combinaciones de secuencias eucarióticas. Dicha competencia permite mantener en ausencia de selección, episomas en la línea humana transformada HEK293A, durante al menos tres meses (80 días) de cultivo continuo.
10. La permisividad a la modificación de los genomas y del contenido en proteínas en los viriones de lentivirus tiene un potencial inmenso de aplicación en transgénesis. Su tolerancia a alteraciones permite anticipar una nueva generación de sistemas genéticos, basados en lentivirus, para modificar el genoma con mínima interferencia.



## **MATERIALES Y MÉTODOS**



## Cultivos celulares

En el presente estudio se utilizaron tres tipos celulares, todos ellos obtenidos de la *American Type Culture Collection* (ATCC; Rockland, MD) y se cultivaron en *Dulbecco's Modified Eagle's Medium* (DMEM) suplementado con 10% de suero fetal bovino (FBS), 1% Glutamax (Invitrogen, Prat del Llobregat, Barcelona, España), 10 mg/ml de antibióticos (penicilina y streptomycin), en condiciones de 37 °C en una atmósfera saturada de humedad (95%) y 5% CO<sub>2</sub>. Las líneas utilizadas fueron:

HEK293: Línea celular humana inmortalizada derivada de células de riñón embrionarias (*Human Embryonic Kidney*, ATCC CRL-1573). Estas células se utilizaron para los experimentos de transfección plasmídica, como prueba de principio y para llevar a cabo las sucesivas reacciones de RMCE.

HEK293T/17 (ATCC CRL-11268): Al igual que la anterior es una línea celular humana proveniente de células de riñón embrionario, pero difiere en el hecho de que fue inmortalizada con el antígeno T grande del virus vacuolizante de simios 40 (en adelante SV40) y sobre la que se escogió el clon 17 debido a su facilidad de transfección. En concreto, esta línea celular se utiliza en el laboratorio (y por extensión en la presente tesis doctoral) de forma rutinaria para la producción de lentivirus recombinantes.

Fibroblastos BJ (ATCC CRL-2522): Línea primaria humana de fibroblastos provenientes de prepucio humano. Son utilizadas debido a su alta capacidad de dividirse, en comparación con otras células primarias humanas, y a la estabilidad de su cariotipo tras sucesivos pases. Se utilizaron para tratar de extrapolar el sistema RMCE a un tipo celular primario y menos permisivo.

A la hora de pasar las células, se lavaron con PBS (Gibco, Invitrogen) y se incubaron 2 minutos a 37° C con tripsina 0.5% (Sigma). Tras lo cual y observando su disgregación, añadimos medio de cultivo completo, para neutralizar dicha reacción y pipetear para disgregar las células y pasarlas a un nuevo soporte. Dichos pases se realizaron cada 3-4 días y fueron de dilución 1/6 en el caso de las células HEK293T, 1/4 en el caso de las 293A y 1/3 para los fibroblastos BJ.

## Generación de plásmidos

Todas las enzimas utilizadas en los clonajes proceden de New England Biolabs y se utilizaron a una concentración final de 5U/microgramo de DNA, durante un tiempo de digestión de entre 3-4 horas y a una temperatura de 37°C. En el caso de las reacciones de ligación, se utilizó la T4 DNA ligasa (New England Biolabs) a una temperatura de 16C y durante un período de 16 horas (toda la noche u *overnight*). Además, todos los plásmidos tras su verificación mediante patrón de digestión en gel de agarosa, también se comprobó su secuencia, a través del servicio de secuenciación del Centro Nacional de Investigaciones Oncológicas (CNIO), mediante el sistema *ABI Prism* (Applied Biosystems) de terminadores fluorescentes *Big Dye™* y el instrumento de electroforesis multicapilar ABI 3730 y su análisis con los programas informáticos FinchTV (Geospizza, PerkinElmer) y/o 4Peaks (Mekentosj).

### Plásmidos KAS

Se utilizaron tres plásmidos diferentes:

1. pKAS, fue utilizado en los experimentos de recombinación transiente y recombinación mediada por copias integradas al azar. Para clonar dicho plásmido, se pidió la síntesis de un bloque (GenScript USA Inc., NJ, USA) que comprende el sitio lox2722, una secuencia stop

previamente descrita (GI:1277147 , <http://www.ncbi.nlm.nih.gov/nuccore/u51223>), el sitio loxP y el gen que codifica para la conferencia de resistencia a higromicina, a este plásmido se le denominó pK0. En dicho plásmido sintetizado, se clonó el promotor de la fosfoglicerato quinasa (en adelante PGK), procedente del plásmido pLVX AcGFP-N1 (Clontech) en los sitios NotI y PstI (plásmido pK1). A continuación se introdujo el cassette de la neomicina (que comprende promotor SV40, el gen de resistencia a neomicina y la señal de poliadenilación de SV40), mediante digestiones XmnI y SalI, tanto en el plásmido pK1, como en el plásmido que lo contiene (pcDNA3.1, Invitrogen), como resultado tuvimos el vector pKAS 1.0.

2. phROSA26, fue utilizado para tratar de generar líneas estables con una única copia integrada de manera controlada y sitio específica. Para su construcción se utilizó un esqueleto vacío derivado del pUC, en concreto, pUC57. En él, se clonaron por PCR los brazos de homología humanos previamente descritos por Keller y colaboradores (ref). Para ello, se utilizó como molde un BAC (*Bacterial Artificial Chromosome*) que los contuviera (RP11-58B17 BAC clone; BACPAC Resource Center (BPRC), Children's Hospital Oakland Research Institute, California, USA). A este plásmido intermedio, le denominamos pUCR26. Para obtener el plásmido final phROSA26, se digirió este plásmido y el plásmido pKAS con las enzimas AscI y MluI, se ligaron ambos fragmentos y se obtuvo el plásmido de recombinación homóloga con la cassette de RMCE contenida en él.
3. pZD KAS2.0, fue utilizado para el mismo fin que el plásmido phROSA26, pero con la salvedad de que este plásmido se utiliza como molde para recombinación homóloga mediada por rotura de la doble hebra de DNA debido a la acción de una nucleasa *ZincFinger*. Se diseñaron unos primers (ver lista o tabla) y se amplificó el promotor híbrido CAG (promotor de la beta actina de pollo combinado con el aplicador de citomegalovirus, CMVenhancer) con extremos NotI y PstI, sustituyéndose en el plásmido pKAS1.0 por el promotor PGK. A continuación, se generó un cassette bicistrónico, en el cual se tuviera la proteína verde fluorescente mejorada (eGFP) y la resistencia a Neo, se llevó a cabo mediante la unión de ambas secuencias de DNA por la introducción de un elemento CHYSEL 2A (*Cis-acting hydrolase element*, [259]) derivado del tescovirus porcino (P2A), evitando la introducción de codones de terminación entre ambos, dado que es necesario que estén en el mismo marco de lectura (ORF, *open reading frame*) y sin ningún tipo de interrupción a excepción del codón de terminación en el extremo 3' terminal. Esto permitirá la generación de un único RNA mensajero, a partir del cual se generarán dos proteínas por la hidrólisis de la secuencia contenida en el elemento de unión o secuencia peptídica 2A, dando lugar a una proteína con una cola adicional de 18 aminoácidos en su extremo C-terminal y una proteína con una prolina en el extremo N-terminal, en ambos casos sin aparente influencia en su funcionalidad [281]. Dicho cassette se amplificó mediante PCR con cebadores específicos que contuvieran las dianas Ddd y Sss en los extremos para su introducción reemplazando el cassette original de resistencia en el plásmido pKAS1.0.

Por último se tomó dicho plásmido y se extrajo la cassette completa para introducirla en el plásmido pZDonor que contiene los brazos de homología para AAVS1, teniendo así la construcción PZD KAS2.0.

### Plásmidos TRINA

- 1) pTRINA, se utilizó en los subsiguientes clonajes de vectores lentivirales lanzadera y para testar la funcionalidad del sistema de manera transitoria. De la misma manera que el plásmido pKAS, la cassette que comprende (en dirección 5—3) lox2722, mCherry ORF, secuencia CHYSEL 2A, loxP fue sintetizada por la empresa GenScript. En dicho plásmido se clonó mediante PCR con extremos compatibles el ORF de la enzima Cre y el promotor del factor de elongación 1 alfa (EF1a).
- 2) LV TRINA, plásmido lanzadera de tercera generación y autoinactivable, derivado de la serie pRRL.sin18.CMV.eGFP.Wpre [102], [203]. Como esqueleto se utilizó el ya mencionado pRRL.sin18 que fue digerido con las enzimas SalI y ClaI al igual que el plásmido pTRINA; siendo clonada la cassette proveniente de este último en el primero.
- 3) LV TRINAdCre, utilizado en los experimentos de optimización de la recombinación. Se generó a partir del vector LVTRINA, en el cual se introdujo un adaptador con la secuencia listada en anexo I (tabla de cebadores); de manera que se reemplazara la secuencia que codifica para la expresión de la enzima Cre desde el promotor EF1a.
- 4) LV Cre, al igual que el anterior fue utilizado en los experimentos de optimización de la reacción de recombinación, en combinación con él. Se generó mediante el clonaje del ORF de Cre mediante el uso de unos cebadores de PCR con extremos que contuvieran sitios de restricción para XbaI y SalI, siendo reemplazado en el vector pRRL.sin18.CMV.eGFP.Wpre en el lugar de la eGFP. De esta manera, quedaría una expresión gobernada por el promotor de CMV.

### Plásmidos Episomales

En este caso, y como en el resto de plásmidos lentivirales,

### Plásmido empaquetador integrasa defectivo

Aunque ampliamente descrito en la literatura y generalizado su uso, se decidió a generar el plásmido pMDLg/pD64VRRE [128]. Para ello, se utilizaron parejas de cebadores que abarcaran la zona de la mutación de la integrasa, conteniendo esta y cebadores externos, tanto en 5' como en el extremo 3'. Se siguió una estrategia de clonaje en la que dichos fragmentos de PCR se introdujeron en un plásmido intermedio para clonaje de fragmentos de PCR (pGEM-T Easy vector, Promega Biolabs) y tras su digestión con las enzimas EcoRV y SalI el fragmento 5', SalI y KpnI el fragmento 3' y EcoRV y KpnI el vector de original sin dicha mutación (pMDLg/pRRE, [102]), se procedió a un llevar a cabo una reacción de ligación con los tres fragmentos para la obtención del plásmido final, manteniendo el correspondiente marco de lectura.

## Transfecciones celulares

### Transfecciones celulares mediante el uso de fosfato de calcio

En el caso de las transfecciones celulares conducentes a la obtención de datos relativos a las reacciones de RMCE y producción de nuevas líneas celulares, se realizaron sobre células HEK293. Para ello, se sembraron, el día previo a la transfección,  $5 \times 10^5$  células en cada pocillo sobre placas de cultivo celular multipocillo de 6 pocillos (BD Biosciences); al día siguiente se observó que su confluencia estuviera entre un 65-80% y con un aspecto saludable. Se les cambió el medio de cultivo por medio fresco y se procedió a preparar la mezcla de transfección.

La transfección se hizo mediante el método de precipitación de fosfato de calcio y DNA [282]; para ello se mezclaron un total de 6 microgramos de DNA plasmídico con 30 microlitros de cloruro de calcio 2.5M en un volumen de 250 microlitros de agua milliQ estéril; por otro lado, se añadieron 250 microlitros de HBSS2x y ambas soluciones se mezclaron, produciendo un burbujeo que favoreciera la formación de los precipitados. Por último, dicha mezcla se añadió sobre un pocillo con células y se devolvieron estas al incubador. Transcurridas 6-8 horas, el medio de transfección fue sustituido por medio fresco previamente atemperado. Transcurridas 24-48 horas se procedió al análisis o subsiguientes tratamientos de dichas células.

### Transfecciones celulares mediante electroporación

En el caso de las electroporaciones en células eucariotas, fueron realizadas sobre los fibroblastos Bj a la hora de tratar de crear la línea celular Bj KAS2.0. Para ello, se mantuvieron las células creciendo en una fase exponencial, sin que llegaran a superar el 95% de confluencia celular (evitando así que se detuviera el crecimiento y fuera más difícil su posterior proliferación). El día de la transfección, se tripsinizaron y se centrifugaron, se les realizó dos lavados con DPBS (Invitrogen, LifeTechnologies), se contaron mediante el uso de una cámara de Neubauer y se resuspendieron en un volumen adecuado de solución R (Invitrogen). Mientras tanto, se preparó una placa de 12 pocillos con medio de cultivo fresco con todos sus componentes, a excepción de los antibióticos (penicilina-estreptomicina), pues su presencia interfiere con la viabilidad celular [283].

A la hora de electroporar las células, se utilizó el sistema *Neon Transfection System*<sup>TM</sup> (Invitrogen, LifeTechnologies) con las puntas de 10 microlitros, para lo que se usó una densidad celular de entre  $5-10 \times 10^7$  células por mililitro; es decir entre 50.000 – 100.000 células por electroporación. Se utilizó el programa previamente optimizado en el que se empleaba un único pulso de 1650 voltios (voltaje) durante 20 milisegundos. Tras ello, las células se ponían en el pocillo y se dejaban en el incubador por 24 horas, momento en el que se analizaba su viabilidad y se procedía al cambio de medio de cultivo. Transcurridas otras 24-48 horas se procedía al uso/análisis de dichas células en subsiguientes ensayos.

## **Lentivirus recombinantes**

### Producción de partículas lentivirales recombinantes

Como se ha comentado en el apartado de plásmidos lentivirales TRINA, los plásmidos utilizados han sido de tercera generación y autoinactivantes; además son vectores pseudotipados [284] con la glicoproteína de la envuelta del virus de la estomatitis vesicular (VSV-G). Para la producción lentiviral se siguió el proceso de producción mediante transfecciones transitorias con precipitados de cloruro de calcio.

El día previo a la transfección se siembran  $11 \times 10^7$  células por placa de 15 centímetros de diámetro (son necesarias o al menos recomendable dos placas para la producción de un sobrenadante lentiviral). Al día siguiente se comprueba que las células estén en una confluencia que va del 70 al 85% y distribuidas en la placa de una manera uniforme y homogénea. Por placa de 15 cm se emplean las siguientes cantidades de cada uno de los plásmidos necesarios para la producción: 6,75 microg del plásmido pRSV-Rev, 8,4 microg del plásmido pMD2.G (VSV-G) y 29,3 microg del plásmido pMDLg/pRRE D64V y 35 microg del plásmido lanzadera utilizado (LV TRINA, LV TRINAdCre, LV Cre, LV iTRINA, , esta mezcla se diluyó en un volumen de 3,6 ml de agua estéril y se añadió 181 microl de Cloruro de calcio 2.5M (Sigma), se mezcló y finalmente se añadieron 3,9 ml de HBS2x. Todo ello se añadió sobre las células gota a gota y sin perturbar demasiado y se devolvió al incubador por un período de 6-8 horas. Transcurrido este tiempo se cambió el medio de transfección por medio de cultivo nuevo y previamente atemperado (18ml por cada placa).

Transcurridas 48 horas el medio fue recogido, se centrifugó en tubos falcon de 50ml (BD Biosceinces) a 1500 rpm durante 5 min. Tras ello se clarificó utilizando filtros de PVDF con baja retención de proteínas con un diámetro de poro de 0.45microm (Steriflip HV Filter Unit, Millipore).

### Concentración de partículas lentivirales recombinantes

El proceso de concentración de las partículas lentivirales recombinantes se lleva a cabo una vez recogidos los sobrenadantes lentivirales, centrifugados y filtrados; en este punto se pueden dividir en las pertinentes alícuotas y ser congeladas a  $-80^{\circ}\text{C}$  o concentrarse, para trabajar con volúmenes menores y poder utilizar un rango de MOIs más amplio. Como protocolo de concentración se utiliza la ultracentrifugación [285], para ello se utilizó una centrífuga Beckman Coulter Optima LX100 con rotor SW28. Se centrifugó a 26.000rpm (90.000g) durante 2 horas a  $4^{\circ}\text{C}$ . El pellet obtenido, se secó del exceso de sobrenadante (con cuidado de no secarlo demasiado) y se resuspendió durante toda la noche en agitación y a  $4^{\circ}\text{C}$  en 500  $\mu\text{l}$  de suero salino 0,9%.

### Titulación de partículas lentivirales recombinantes

A la hora de titular las partículas lentivirales se pueden utilizar diferentes métodos (qPCR, FACS, ELISA, número de colonias resistentes,...) [286]; de rutina en el laboratorio se utilizan tanto la fluorescencia por FACS (unidades transductoras por ml) y/o el número de partículas obtenido mediante qPCR, dependiendo de si el transgen lleva sistema reportero o no, respectivamente.

- 1) FACS. Como se ha comentado se utilizó en los casos que se disponía de un gen chivato o reportero; para ello se utilizó la transducción de los virus obtenidos sobre células HEK 293T/17. Al igual que en el caso de las transfecciones, las células se sembraron 24 h antes, pero en este caso en placas multipocillo de 24 pocillos, a una densidad de 60.000 células/ $\text{cm}^2$ , es decir, 120.000 células por cada pocillo. Transcurridas 24 horas, se comprobó el estado de las células y se procedió a su transducción, para ello se utilizaron diluciones de los concentrados

lentivirales en medio de cultivo fresco, al cual se le añadió polibreno (Sigma) a una concentración final de 8 microg/ml. Las diluciones utilizadas para determinar el título van desde 1/50 hasta 1/5.000. Transcurridas 6 horas el medio de inoculación fue reemplazado por medio fresco y a las 72 horas se tripsinizaron las células para su recolección y determinación del porcentaje de fluorescencia mediante el análisis por FACS.

Una vez obtenido el dato, se aplicó la fórmula para determinar el número de unidades transductoras, para ello:

[Número de células a día 0 (24 h tras la siembra) x % células positivas FACS x factor de dilución] / volumen del virus añadido (ml)

Ello nos dará las partículas infectivas o transductoras por ml, los valores oscilan entre  $10^7$ - $10^8$ .

- 2) PCR cuantitativa en tiempo real (*Real Time* qPCR). Se usó en los casos en que el vector lanzadera utilizado no poseyera codificación alguna para proteína de color detectable por citometría de flujo. Como molde para la reacción de detección se midió la señal que daval a secuencia stop del cDNA (DNA complementario) de la cadena negativa, que está presente en todas las cápsidas lentivirales [287], sin necesidad de llevar a cabo retrotranscripción.

La reacción se llevó a cabo con los valores que se detallan en la siguiente tabla:

| Reactivo                        | Concentración inicial | Concentración final | Volumen en $\mu$ l |
|---------------------------------|-----------------------|---------------------|--------------------|
| <i>TaqMan Master Mix</i> 2x     | 2x                    | 1x                  | 12,5               |
| Sobrenadante viral *            | -                     | -                   | 1                  |
| Nucleótidos trifosfatos (dNTPs) | 5 mM                  | 250 $\mu$ M         |                    |
| LTR Fw                          | 10 mM                 | 600 $\mu$ M         |                    |
| LTR Rv                          | 10 mM                 | 600 $\mu$ M         |                    |
| Sonda LTR                       | 10 $\mu$ M            | 200 nM              |                    |
| Agua mQ                         | -                     | -                   | Hasta 25           |

(\*) con distintas diluciones seriadas

En primer lugar se deberá diseñar una recta patrón que nos permita interpolar nuestros datos (a posteriori) conforme a un patrón previamente testado y validado, que nos permita establecer un límite de detección y un rango lineal; para ello, se hicieron diluciones seriadas del plásmido pRRL.sin18.CMV.eGFP.Wpre, con un número de copias conocido, que abarca desde  $10^9$  hasta  $10^2$ , ello se sometió a Real Time qPCR por duplicado, en el equipo *ABI PRISM® 7900HT Sequence Detection System*. Utilizando el programa que viene diseñado por defecto, 2 min a 50°C, desnaturalización a 95°C durante 10 min, seguido de 40 ciclos con un programa de PCR en dos pasos (desnaturalización e hibridación/elongación) que son desnaturalización 15 segundos a 95°C e hibridación/elongación 1 min a 60°C.

El análisis de los resultados se llevó a cabo utilizando en el software SDS v2.3 (Sequence Detection System, Applied Biosystem) y extrayendo las Ct, se interpolaron en la parte lineal de la curva que nos ofrecía el patrón previamente generado. Se ha de tener en cuenta que los títulos obtenidos mediante este método son dos órdenes de magnitud a los obtenidos mediante citometría de flujo, por lo que únicamente el 1% de las partículas detectadas son infectivas [288] y dicha corrección se aplicó en los resultados obtenidos



### Selección de clones celulares

En el proceso de obtención tanto de las células parentales KAS, como de los posibles productos celulares de la recombinación (RMCE) TRINA, las células fueron sometidas a presión selectiva para de esta manera obtener clones aislados. El proceso llevado a cabo fue el mismo, una vez las células fueron sometidas a transfección o a electroporación (KAS) o tras la transducción con los distintos IDLV (TRINA).

Se mantuvieron las células durante 72 horas en medio no selectivo (sin antibiótico), para de esta manera permitir la mayor viabilidad celular posible tras el proceso de introducción de material exógeno (transfección, electroporación); transcurrido dicho período se las sometió a la presión selectiva; esto es, se añadió Higromicina (Sigma) o G418 (Geneticina, Invitrogen) al medio de cultivo en concentraciones determinadas por una curva de muerte celular, lo cual estableció los valores para los fibroblastos humanos (Bj) en 12,5 µg/ml y 150 µg/ml (higromicina y geneticina, respectivamente) y para la línea celular HEK293A en 50 µg/ml y 450 µg/ml, respectivamente. El medio de cultivo con el antibiótico se mantuvo con cambios cada 2-3 días durante un período de 8 días en el caso de la higromicina y de 14 días para la geneticina. Una vez los clones aparecieron y siempre estando lo suficientemente aislados en la placa de cultivo, fueron seleccionados y picados manualmente bajo una lupa estereoscópica y posteriormente sembrados para su amplificación; por otro lado, y en un experimento paralelo (dúplica) se procedió a su cuantificación tras ser teñidos con cristal violeta.

En el caso de la cuantificación se realizó mediante una tinción con cristal violeta y su posterior conteo. Para lo cual se colocaron las placas a teñir en hielo y se lavaron 2 veces con PBS estéril, se fijaron con metanol 100% a -20°C durante 10 minutos, se retiró el metanol y las placas se pusieron a temperatura ambiente y sin permitir que se secaran por completo se añadió la solución de tinción (cristal violeta 0,5% en metanol al 25%) para cubrir el fondo de la placa. Se incubó a temperatura ambiente y con una leve agitación entre 7-15 minutos, tras lo cual se retiró dicha solución y se lavó por duplicado (con agua destilada) hasta quitar el exceso de solución de tinción. Por último se invirtieron las placas para secarlas y se escanearon o contabilizaron el número de colonias.

### Caracterización genética de clones celulares

A la hora de realizar la caracterización genética de los clones aislados, se utilizaron distintas técnicas (PCR, qPCR, Southern Blot y FISH), exceptuando el caso de la FISH (que requiere de un protocolo específico) tanto para PCR, qPCR, como para el Southern Blot, se requería de un aislamiento previo del material genético (DNA) que se pretende caracterizar.

#### Extracción y purificación del DNA genómico

En primer lugar se amplificaron las células para disponer de un número suficiente como para poder tener el material necesario para caracterizarlas. Tras disponer de un número considerable de células (entre  $5 \times 10^5$  y  $1 \times 10^7$ ), se tripsinizaron y recogieron, a continuación se centrifugaron 5 minutos a 1250rpm y se lavaron con 2 ml de PBS, tras lo cual se volvieron a centrifugar (en este punto pueden ser congelados los pellet a -20°C o continuar con el proceso de extracción). Se resuspendieron en 250 µl del buffer de digestión, se incubaron las muestras en agitación toda la noche a 50°C, una vez transcurrido se añadieron 250 µl de fenol:cloroformo:alcohol isoamílico (Sigma, proporción 25:24:1) y se mezcló y centrifugó a 12000 rpm 10 min. Se tomó la fase acuosa superior y se descartó el resto, se transfirió a un nuevo eppendorf y se añadió cloroformo, se repitió el proceso. Una vez limpio de fenol, nos quedamos con la fase acuosa y se procede a la precipitación del DNA, para ello

se añaden 25 µl de acetato sódico 3M y 500 µl de etanol absoluto (Fluka). Se centrifuga 10 min a 4°C y 12.000rpm. Se descarta el sobrenadante, y se lava el pellet con etanol 70%, se centrifuga 5 min a 12.000 rpm y 4°C se desecha el sobrenadante y se seca el pellet (sin excedernos para una adecuada resuspensión). Por último, se resuspende en buffer TE y se guarda a -20°C, para su posterior utilización.

#### PCR de recombinación homóloga e integración específica

Para comprobar si la inserción de la cassette KAS, el reemplazamiento de TRINA y/o la inserción del Sistema Reportero de Splicing (SRS) ocurrió sitio específica, necesitamos de la utilización de distintas parejas de primers que hibriden tanto en el interior de dichas cassettes, como parejas de primers que lo hagan tanto en el interior, como en las regiones que flanquean el lugar de inserción tanto en 5' como en 3'. Para ello se usaron:

- 1) Detección de integración: Primers eGFP Fw y NeoRv (KAS y SRS)
- 2) Región de homología 5': AAVS1 LA y CMV Rv o PGK Rv (dependiendo de la construcción KAS utilizada). CMV Fw y Cherry Rv (Reemplazamiento TRINA). AAVS1 Fw y CMV Rv (SRS)
- 3) Región de homología 3': AAVS1 RA y Hygro Fw (KAS). EF1a e Hygro Rv (Reemplazamiento TRINA). AAVS1 Rv y Neo Fw (SRS).

La polimerasa que se utilizó es la Phusion High-Fidelity DNA Polymerase (New England Biolabs) y el protocolo es:

| Reactivo             | Concentración inicial | Concentración final | Volumen en µl |
|----------------------|-----------------------|---------------------|---------------|
| <i>5x Buffer GC</i>  | 5x                    | 1x                  | 4             |
| dNTPs                | 10mM                  | 200 µM              | 1             |
| Forward primer       | 10 µM                 | 0.5 µM              | 1             |
| Reverse primer       | 10 µM                 | 0.5 µM              | 1             |
| DNA molde (genómico) | x                     | 2.5 ng/ µl          | 2 (máximo)    |
| DMSO                 | 100%                  | 2%                  | 0.4           |
| Agua mQ              | -                     | -                   | Hasta 20      |

El programa utilizado es, desnaturalización a 98°C durante 1 min, seguido de 35 ciclos con un programa de PCR en tres pasos (desnaturalización, hibridación y elongación) que son desnaturalización 8 segundos a 98°C, hibridación 20 segundos a 61°C y elongación 1 minuto/2Kb a 72°C; por último se lleva a cabo un ciclo final de elongación durante 10 minutos a 72°C y se pasa inmediatamente al hielo (4°C) hasta su posterior uso.

#### Southern Blot análisis

Se utilizó en los experimentos para la determinación de la presencia de integraciones (o ausencia) en la elección de un sistema episomal óptimo. Como sonda se utilizó tanto una región perteneciente a las LTRs, como 500 pb del gen que codifica para la eGFP. La manera en que se hizo se describe a continuación:

Se digirieron 10 µg de DNA toda la noche. A continuación se cargaron las muestras en un gel de agarosa al 0.8% en TAE y se tuvo a un voltaje de 25V toda la noche. Se le realizó una fotografía al gel y se procedió al proceso de transferencia. Se despurinizó el DNA del gel con ácido clorhídrico (Merck) 0.25M durante 15 minutos, se neutralizó y desnaturalizó el gel con hidróxido sódico (Merck) 0.4M durante 45 minutos. Mientras tanto, se preparó la membrana (Hybond N+ , Amersham GE) hidratándola en agua destilada y tratándola posteriormente con hidróxido sódico. Se realiza la transferencia por un sistema de capilaridad y peso durante toda la noche.

Al día siguiente se recoge la membrana y se hace fija el DNA a la membrana mediante un tratamiento crosslinking (Stratalinker UV crosslinker, Stratagene). Se hidrata la membrana en buffer 2xSSC y se prehibrida a 56°C con la solución de prehibridación y DNA de esperma de salmón (previamente desnaturalizado); durante 3 horas en agitación.

Mientras se marca la sonda (Amersham rediprime II random prime labelling, Amersham GE) con fósforo radiactivo ( $P^{32}$ ) y una vez desnaturalizada se añade a la membrana previamente prehibridada. Se mantiene en agitación toda la noche a 65°C.

Al día siguiente se lava la membrana, 2xSSC 5 minutos, 1xSSC/0.1%SDS 20 minutos a 65°C (dos veces), 0.2xSSC/0.1%SDS 20 minutos a 65°C. Último lavado de 2xSSC antes de montaje del sistema para su revelado. Se reveló mediante el sistema Phosphorimager (Storm 860 Molecular Imager, Amersham GE).

#### Cuantificación de número de copias por análisis mediante PCR cuantitativa.

El DNA genómico se extrajo de la misma manera que se ha descrito con anterioridad; una vez cuantificado, tanto por comparación con un marcador de pesos moleculares (1Kbp marker, Plasmid Factory) como utilizando diluciones seriadas en Nanodrop ND1000 (NanoDrop Technologies). La PCR se realizó mediante metodología de SyBR Green (Applied biosystems). El número de copias integradas, tanto en los casos de las células KAS, TRINA, como en la detección de copias episomales, se realizó conforme a lo descrito previamente por Butler y colaboradores [122] con algunas modificaciones. Las curvas standard se obtuvieron por diluciones que van desde 0.01 pg hasta 100ng, lo que correspondería a un número de copias desde  $10^2$  hasta  $10^9$ .

Los valores de Ct obtenidos fueron interpolados en las correspondientes curvas y el número de copias relativo calculado. El número de equivalentes celulares fue calculado usando valores de Ct de un gen de copia única, como es la Albumina, en muestras de DNA genómico.

La reacción de qPCR siguió las siguientes condiciones:

El programa utilizado es, desnaturalización a 95°C durante 10 min, seguido de 40 ciclos con un programa de PCR en tres pasos (desnaturalización, hibridación y elongación) que son desnaturalización 15 segundos a 95°C, hibridación 20 segundos a 60°C y elongación 30 segundos a 72°C; por último se lleva a cabo un ciclo final de elongación durante 10 minutos a 72°C y se hace un análisis de melting, para ver la eficiencia y especificidad de los primers.

#### Análisis Citogenético. FISH y Cariotipado

Tanto la determinación del cariotipo de las células HEK293A, como su análisis de localización de episomas, integraciones cromosómicas (o ausencia de ellas), como el *site specific targeting*, fue corroborado mediante el uso de alguna de estas técnicas.

##### 1. Citogenética Convencional (Cariotipos de Bandas G)

El análisis citogenético convencional consiste en el estudio morfológico de los cromosomas metafásicos teñidos y bandeados fundamentalmente con bandas G (tripsina y giemsa). Esta técnica permite el análisis de posibles alteraciones cromosómicas, tanto numéricas como estructurales.

###### (i) Preparación de los cromosomas

El estudio se lleva a cabo tras el cultivo en suspensión de células de sangre periférica. El medio de cultivo que empleamos es el RPMI suplementado con 2% de fitohemaglutinina, 50 mg/ml de

gentamicina, 10% de suero bovino fetal, 1.5% de glutamina, 1.5% de Hepes, antibióticos (penicilina y estreptomycin) y fungizona. Para el cultivo se utilizan 5 ml de medio suplementado y 1 ml de sangre periférica, se incuban durante 72 horas en estufa a 37° C. Posteriormente, se añade un 0.1% del volumen de colchicina, lo que provoca el paro de las células en división en el estado de metafase. Para romper la membrana celular se tratan las células con solución hipotónica KCl 0.07 M durante media hora en estufa a 37° C, tras esto se fijan (metanol:ácido acético, 3:1) y se hacen extensiones sobre portaobjetos.

#### (ii) Obtención bandas G

Envejecer la extensión durante 10 minutos en placa térmica a 90° C. Tratar las extensiones con una solución de tripsina (200 ml solución tampón fosfato; 800 mg ClNa; 20 mg Tritiplex; 50 mg Tripsina 1:250) a 37° C durante unos 10 segundos (tiempo variable). Para neutralizar la acción de la tripsina se lavan los portas con buffer fosfato a pH 6.8. Teñir con Leishman y buffer fosfato (1:3) durante 2 minutos. Lavar con agua destilada.

#### (iii) Captura de imágenes

La captura de imágenes a través de un programa informático (CytoVision System) se realiza mediante un microscopio (Olympus AX60) acoplado a una cámara digital (Cohu). Empleamos el mismo programa para la generación de los cariotipos.

### 2. Hibridación *in situ* fluorescente (FISH)

La técnica de hibridación *in situ* fluorescente (FISH) se basa en la hibridación de una sonda de ADN marcada con un fluorocromo sobre su secuencia complementaria del genoma. Permite detectar y localizar secuencias específicas sobre preparaciones cromosómicas, extensiones celulares y cortes de tejidos frescos o incluidos en parafina. Esta técnica se usa desde los años 80 [289]. Se pueden utilizar de forma general tres tipos de sondas: centroméricas, de pintado cromosómico y de secuencia única. Las primeras están compuestas por una secuencia de ADN repetitivo que hibrida en la región centromérica del cromosoma. Permiten detectar alteraciones numéricas en metafases y núcleos interfásicos. Por otro lado, las sondas de pintado cromosómico están compuestas por una serie de sondas que hibridan a lo largo de todo un cromosoma. Permiten detectar alteraciones numéricas y estructurales sobre metafases. Por último, las sondas de secuencia única hibridan con el ADN de una región genómica concreta (gen o banda cromosómica) en metafases y núcleos interfásicos. El protocolo que seguimos fue el recomendado por la casa comercial con modificaciones:

#### (i) Preparación de las sondas

Un ml de cada sonda se mezcla con 7 ml de solución ó buffer de hibridación comercial LSI/WCP (que contiene dextrán sulfato, formamida y SSC (pH 7.0)) (Vysis, Downers Grove, USA) y 3 ml de agua destilada. Se deja esta mezcla de hibridación unos 30 minutos en estufa a 37° C para pre-anillamiento.

#### (ii) Preparación de la muestra

Las preparaciones se envejecen colocándolas en una placa térmica a 90° C durante 10 minutos. Posteriormente se deshidratan las muestras pasándolos los portas por una serie de alcoholes de gradación creciente (70%, 80% y 100%). Por último, se dejan secar los portas a temperatura ambiente.

### (iii) Hibridación e incubación

Se añade la mezcla de hibridación sobre una zona del porta seleccionada, y se coloca un cubreobjetos. Se desnaturalizan sonda y muestra en un único paso, en placa térmica a 75° C durante un minuto. Se incuba toda la noche a 37° C (en cámara húmeda).

### (iv) Lavados

Primero, se pasan los portas por una solución de 0.4xSSC con 0.3% de NP-40 durante dos minutos a 75° C. Tras esto, se dejan durante 5 minutos en agitación a temperatura ambiente en 2xSSC con 0.1% de NP-40. Finalmente, se lavan en PBD 1x a temperatura ambiente dos minutos más. Se deja escurrir el exceso de PBD y se contratiñen con una mezcla 1:1 de DAPI (Vysis, Downers Grove, USA) y antifade (Qbiogene).

### (v) Captura y análisis de imágenes

Para la captura de imágenes utilizamos un microscopio epifluorescente (Olympus AX60) acoplado a una cámara digital (Sensys). El programa informático que utilizamos para la captura y posterior análisis de imágenes fue Cytovision System.

## 3. Obtención de sondas para FISH

El proyecto de secuenciación del genoma humano desarrollado por varios centros públicos ha generado una enorme fuente de librerías de clones mapeados y secuenciados que están revolucionando la citogenética molecular. Ahora, cualquier región de cualquier cromosoma puede ser estudiada en Internet usando uno de los buscadores genómicos como el del Ensembl ([http://www.ensembl.org/Homo\\_sapiens/cytoview](http://www.ensembl.org/Homo_sapiens/cytoview)), el desarrollado por la Universidad de Santa Cruz de California (UCSC) (<http://genome.cse.ucsc.edu/index.html>), o el del National Center for Biotechnology Information (<http://www.ncbi.nlm.nih.gov/mapview/>). Los clones se pueden seleccionar y adquirir fácilmente. En los últimos años, los Cromosomas Artificiales de Bacterias (BACs) se han confirmado como el sistema vector elegido para la construcción de estas librerías de ADN. El sistema de clonaje de BACs permite el aislamiento de fragmentos de ADN genómico de entre 100-200 Kb, lo suficientemente grandes como para poder usarlos como sondas en estudios de FISH.

### (i) Sondas de secuencia única

Utilizamos las bases de datos Ensembl Cytoview para seleccionar los BACs que contengan genes o regiones génicas de interés. Los clones se adquieren en el “Children Hospital Oakland Research Institute” (CHORI) (<http://www.chori.org/bacpac/>) y en el CalTech BAC (<http://informa.bio.caltech.edu/>). Los clones se reciben en medio LB con 20 mg/ml de cloranfenicol. Como estos cultivos tienen una vida finita, es recomendable preparar soluciones stock de glicerol para poder mantener indefinidamente las bacterias congeladas a -80° C.

### (ii) Marcaje de la sonda

El ADN de los BACs se marcó fluorescentemente para poder usarlo como sonda en experimentos de FISH. El marcaje se hizo con el método denominado “nick translation”. Este tipo de marcaje es catalizado por dos enzimas, la ADNasa I y la ADN polimerasa I. La primera enzima, en presencia de iones magnesio, se convierte en una endonucleasa de cadena sencilla, y se añade para que genere cortes (“nicks”) al azar en las dos hebras de cualquier molécula de ADN. La ADN polimerasa I de *Escherichia coli*, mediante su actividad exonucleasa 5'→3' elimina nucleótidos del extremo 5', al mismo tiempo que por su actividad polimerasa 5'→3' añade nucleótidos (marcados) a todos los extremos 3'OH disponibles creados por la ADNasa, copiando la secuencia de la cadena complementaria. Esta actividad exonucleasa/polimerasa, mueve (“translate”) cualquier corte de

cadena sencilla en la dirección 5'→3'. Cuando se encuentran cortes de las dos hebras, la molécula de ADN se rompe. Como los cortes son generados al azar, el ADN se marca uniformemente.

## Caracterización fenotípica

### Immunofluorescencia

Para la realización de inmunofluorescencia de células transfectadas y/o infectadas con plásmidos que expresen la enzima cre recombinasa, se sembraron las células en cristales tratados con poli-L-lisina (Sigma), aproximadamente 50.000 células/pocillo M24. Transcurridas 24-48 horas, se lavaron las células con PBS y se fijaron durante 10 minutos con Paraformaldehído al 4% (USB, Affymetrix) en PBS; se permeabilizaron utilizando 0.5% TritonX-100 (Sigma) y PBS durante 10 minutos a temperatura ambiente y se bloqueó con suero de cabra al 4% (Sigma) en PBS. En dicho buffer se realizaron las diluciones de los anticuerpos, utilizando una dilución 1/200 del anticuerpo anti-Cre (Novagen EMD4Biosciences, Millipore) durante 1 hora a temperatura ambiente.

Como anticuerpos secundarios se utilizaron distintas opciones dependiendo de la disponibilidad de los mismos en la unidad de microscopía del centro, entre ellos, cabra anti-conejo Alexa-633 conjugado o cabra anti-ratón Alexa-633 conjugado (Molecular probes, Invitrogen). Por último, las muestras se tiñeron con DAPI Gold antifade (Molecular probes, Invitrogen), que sirvió de medio de montaje y para teñir el DNA presente en los núcleos.

Las muestras fueron visualizadas en microscopio confocal Leica SPE (Leica microsystems) o en Zeiss LSM 700 (Carl Zeiss), usando 4 lasers que serían 405nm (DAPI), 488nm (eGFP), 564nm (mCherry) y 633nm (anticuerpo secundario). Los datos fueron adquiridos de manera secuencial a una resolución de 1024x1024 pixels y correspondientes a secciones de un grosor no mayor a 1 µm.

Finalmente las imágenes fueron tratadas con los paquetes de software Adobe Photoshop (Adobe Systems Incorporated) y/o ImageJ (NIH, USA)

### Western Blot

Para ello se llevó a cabo métodos standard de transferencia de proteínas a membranas de PVDF, con la variación de que dicha transferencia se hizo utilizando el sistema iBlot (Invitrogen). Las membranas se marcaron para la detección de Cre (Novagen EMD4Biosciences, Millipore) a una dilución 1/1000 o tubulina (AbCam, ab6046) a dilución 1/2000, ambos en PBS Tween al 0.1% (Merck). Como anticuerpo secundario se utilizó un anticuerpo contra ratón hecho en cabra contra la región IgG (Santa Cruz Biotechnology) y se reveló con el kit ECL Plus (GE Helthcare) exponiéndolo a una membrana en una reveladora Kodak Biomax (Kodak).

### FACS

El análisis de la expresión de células fluorescentes TRINA, se cuantificó finamente por la metodología de citometría de flujo; para ello, se recogieron las células (tripsinización), se lavaron con PBS y se dispusieron en tubos de citómetro para su análisis en el equipo de citómetro de flujo FACS Aria Cell Sorter (BD Biosciences) utilizando un fotomultiplicador con láser 561nm (en el caso de mCherry) y otro con filtro 488nm (para detección de eGFP).

## **BIBLIOGRAFÍA**





- [1] J. A. Wolff and J. Lederberg, "An early history of gene transfer and therapy," *Hum. Gene Ther.*, vol. 5, no. 4, pp. 469–480, Apr. 1994.
- [2] O. T. Avery, C. M. MacLeod, and M. McCarty, "STUDIES ON THE CHEMICAL NATURE OF THE SUBSTANCE INDUCING TRANSFORMATION OF PNEUMOCOCCAL TYPES : INDUCTION OF TRANSFORMATION BY A DESOXYRIBONUCLEIC ACID FRACTION ISOLATED FROM PNEUMOCOCCUS TYPE III.," *J. Exp. Med.*, vol. 79, no. 2, pp. 137–158, Feb. 1944.
- [3] J. D. Watson and F. Crick, "nature physics portal - looking back - DNA - a new twist on life," *Nature*, 1953.
- [4] T. Edmunds, S. M. Van Patten, J. Pollock, E. Hanson, R. Bernasconi, E. Higgins, P. Manavalan, C. Ziomek, H. Meade, J. M. McPherson, and E. S. Cole, "Transgenically produced human antithrombin: structural and functional comparison to human plasma-derived antithrombin," *Blood*, vol. 91, no. 12, pp. 4561–4571, Jun. 1998.
- [5] D. M. Stalker, K. E. McBride, and L. D. Malyj, "Herbicide resistance in transgenic plants expressing a bacterial detoxification gene," *Science*, vol. 242, no. 4877, pp. 419–423, Oct. 1988.
- [6] T. H. Schuler, G. M. Poppy, B. R. Kerry, and I. Denholm, "Insect-resistant transgenic plants," *Trends in Biotechnology*, 1998.
- [7] A. M. Gualandi-Signorini and G. Giorgi, "Insulin formulations-a review," ... *review for medical and ...*, 2001.
- [8] T. Von Fange and T. McDiarmid, "Can recombinant growth hormone effectively treat idiopathic short stature?," *Journal of family ...*, 2008.
- [9] K. R. Thomas and M. R. Capecchi, "Site-directed mutagenesis by gene targeting in mouse embryo-derived stem cells," *Cell*, vol. 51, no. 3, pp. 503–512, Nov. 1987.
- [10] D. Hanahan, E. F. Wagner, and R. D. Palmiter, "The origins of oncomice: a history of the first transgenic mice genetically engineered to develop cancer," *Genes Dev.*, vol. 21, no. 18, pp. 2258–2270, Sep. 2007.
- [11] C. N. Stewart Jr, "Go with the glow: fluorescent proteins to light transgenic organisms," *Trends in Biotechnology*, vol. 24, no. 4, pp. 155–162, Apr. 2006.
- [12] P. Berg, "Meetings that changed the world: Asilomar 1975: DNA modification secured," *Nature*, vol. 455, no. 7211, pp. 290–291, Sep. 2008.
- [13] T. Friedmann, "A brief history of gene therapy," *Nature Genetics*, vol. 2, no. 2, pp. 93–98, Oct. 1992.
- [14] T. Friedmann and R. Roblin, "Gene Therapy for Human Genetic Disease?," *Science*, vol. 175, no. 4025, pp. 949–955, Mar. 1972.
- [15] S. Rogers and P. Pfuderer, "Use of viruses as carriers of added genetic information," *Nature*, vol. 219, no. 5155, pp. 749–751, Aug. 1968.
- [16] S. Rogers, A. Lowenthal, H. G. Terheggen, and J. P. Columbo, "Induction of arginase activity with the Shope papilloma virus in tissue culture cells from an argininemic patient," *J. Exp. Med.*, vol. 137, no. 4, pp. 1091–1096, Apr. 1973.
- [17] H. G. Terheggen, A. Lowenthal, F. Lavinha, J. P. Colombo, and S. Rogers, "Unsuccessful trial of gene replacement in arginase deficiency," *Z Kinderheilkd*, vol. 119, no. 1, pp. 1–3, 1975.
- [18] M. J. Cline, H. Stang, K. Mercola, L. Morse, and R. Ruprecht, "Gene transfer in intact animals. - Abstract - UK PubMed Central," *Nature*, 1980.
- [19] S. Hacein-Bey-Abina, C. von Kalle, M. Schmidt, F. Le Deist, N. Wulffraat, E. McIntyre, I. Radford, J.-L. Villeval, C. C. Fraser, M. Cavazzana-Calvo, and A. Fischer, "A serious adverse event after successful gene therapy for X-linked severe combined immunodeficiency," *N. Engl. J. Med.*, vol. 348, no. 3, pp. 255–256, Jan. 2003.
- [20] E. Marshall, "Gene therapy death prompts review of adenovirus vector," *Science*, vol. 286, no. 5448, pp. 2244–2245, Dec. 1999.
- [21] J. A. Thomson, J. Itskovitz-Eldor, S. S. Shapiro, and M. A. Waknitz, "Embryonic stem cell lines derived from human blastocysts," *Science*, 1998.
- [22] K. Takahashi and S. Yamanaka, "Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors," *Cell*, vol. 126, no. 4, pp. 663–676, Aug. 2006.

- [23] F. Soldner, J. Laganière, A. W. Cheng, D. Hockemeyer, Q. Gao, R. Alagappan, V. Khurana, L. I. Golbe, R. H. Myers, S. Lindquist, L. Zhang, D. Guschin, L. K. Fong, B. J. Vu, X. Meng, F. D. Urnov, E. J. Rebar, P. D. Gregory, H. S. Zhang, and R. Jaenisch, "Generation of isogenic pluripotent stem cells differing exclusively at two early onset Parkinson point mutations.," *Cell*, vol. 146, no. 2, pp. 318–331, Jul. 2011.
- [24] M. M. Gottesman, "Cancer gene therapy: an awkward adolescence," *Cancer Gene Ther*, vol. 10, no. 7, pp. 501–508, Jul. 2003.
- [25] X. Gao, K.-S. Kim, and D. Liu, "Nonviral gene delivery: what we know and what is next.," *AAPS J*, vol. 9, no. 1, pp. E92–104, 2007.
- [26] U. Lakshmipathy, B. Pelacho, K. Sudo, J. L. Linehan, E. Coucouvanis, D. S. Kaufman, and C. M. Verfaillie, "Efficient transfection of embryonic and adult stem cells.," *Stem Cells*, vol. 22, no. 4, pp. 531–543, 2004.
- [27] H. Siemen, M. Nix, E. Endl, P. Koch, J. Itskovitz-Eldor, and O. Brüstle, "Nucleofection of human embryonic stem cells.," *Stem Cells Dev.*, vol. 14, no. 4, pp. 378–383, Aug. 2005.
- [28] L. Vallier, P. J. Rugg-Gunn, I. A. Bouhon, F. K. Andersson, A. J. Sadler, and R. A. Pedersen, "Enhancing and diminishing gene function in human embryonic stem cells.," *Stem Cells*, vol. 22, no. 1, pp. 2–11, 2004.
- [29] D. Anderson, T. Self, I. R. Mellor, G. Goh, S. J. Hill, and C. Denning, "Transgenic enrichment of cardiomyocytes from human embryonic stem cells.," *Mol. Ther.*, vol. 15, no. 11, pp. 2027–2036, Nov. 2007.
- [30] M. A. Kay, "State-of-the-art gene-based therapies: the road ahead.," *Nat Rev Genet*, vol. 12, no. 5, pp. 316–328, May 2011.
- [31] J. J. Skehel and D. C. Wiley, "Receptor binding and membrane fusion in virus entry: the influenza hemagglutinin," *Annual review of biochemistry*, 2000.
- [32] C. A. Benedict, P. S. Norris, and C. F. Ware, "To kill or be killed: viral evasion of apoptosis," *Nat. Immunol.*, 2002.
- [33] M. A. Kay, J. C. Glorioso, and L. Naldini, "Viral vectors for gene therapy: the art of turning infectious agents into vehicles of therapeutics.," *Nat. Med.*, vol. 7, no. 1, pp. 33–40, Jan. 2001.
- [34] V. R. Racaniello and D. Baltimore, "Cloned poliovirus complementary DNA is infectious in mammalian cells.," *Science*, vol. 214, no. 4523, pp. 916–919, Nov. 1981.
- [35] C. E. Thomas, A. Ehrhardt, and M. A. Kay, "Progress and problems with the use of viral vectors for gene therapy," *Nat Rev Genet*, vol. 4, no. 5, pp. 346–358, May 2003.
- [36] R. Waehler, S. J. Russell, and D. T. Curiel, "Engineering targeted viral vectors for gene therapy.," *Nat Rev Genet*, vol. 8, no. 8, pp. 573–587, Aug. 2007.
- [37] R. M. Eager and J. Nemunaitis, "Clinical development directions in oncolytic viral therapy.," *Cancer Gene Ther*, vol. 18, no. 5, pp. 305–317, May 2011.
- [38] D. Stone and A. Lieber, "New serotypes of adenoviral vectors.," *Curr. Opin. Mol. Ther.*, 2006.
- [39] C. E. Gomez, J. L. Najera, and M. Krupa, "The poxvirus vectors MVA and NYVAC as gene delivery systems for vaccination against infectious diseases and cancer," *Current gene ...*, 2008.
- [40] B. K. Berges, J. H. Wolfe, and N. W. Fraser, "Transduction of brain by herpes simplex virus vectors.," *Mol. Ther.*, vol. 15, no. 1, pp. 20–29, Jan. 2007.
- [41] D. Hüser, A. Gogol-Döring, T. Lutter, S. Weger, K. Winter, E.-M. Hammer, T. Cathomen, K. Reinert, and R. Heilbronn, "Integration preferences of wildtype AAV-2 for consensus rep-binding sites at numerous loci in the human genome.," *PLoS Pathog.*, vol. 6, no. 7, p. e1000985, 2010.
- [42] A. Schröder, P. Shinn, H. Chen, C. Berry, and J. R. Ecker, "HIV-1 integration in the human genome favors active genes and local hotspots," *Cell*, 2002.
- [43] P. Hasty, J. Rivera-Pérez, and A. Bradley, "The length of homology required for gene targeting in embryonic stem cells.," *Mol. Cell Biol.*, vol. 11, no. 11, pp. 5586–5591, Nov. 1991.
- [44] P. Sung and H. Klein, "Mechanism of homologous recombination: mediators and helicases take on regulatory functions.," *Nat. Rev. Mol. Cell Biol.*, vol. 7, no. 10, pp. 739–750, Oct. 2006.
- [45] A. Colosimo, K. K. Goncz, A. R. Holmes, K. Kunzelmann, G. Novelli, R. W. Malone, M. J. Bennett, and D. C. Gruenert, "Transfer and expression of foreign genes in mammalian cells.," *BioTechniques*, vol. 29, no. 2, pp. 314–8– 320–2– 324 passim, Aug. 2000.
- [46] K. M. Vasquez, K. Marburger, and Z. Intody, "Manipulating the mammalian genome by homologous recombination," presented at the Proceedings of the ..., 2001.

- [47] K. Smith, "Theoretical mechanisms in targeted and random integration of transgene DNA.," *Reprod. Nutr. Dev.*, vol. 41, no. 6, pp. 465–485, Nov. 2001.
- [48] P. E. Monahan and R. J. Samulski, "AAV vectors: is clinical success on the horizon?," *Gene Ther.*, 2000.
- [49] R. J. Yáñez and A. C. Porter, "Therapeutic gene targeting.," *Gene Ther.*, vol. 5, no. 2, pp. 149–159, Feb. 1998.
- [50] N. Muzyczka and K. I. Berns, *Muzyczka: Parvoviridae: the viruses and their replication - Google Scholar*. Fields virology, 2001.
- [51] M. A. F. V. Gonçalves, "Adeno-associated virus: from defective virus to effective vector.," *Virology*, vol. 2, p. 43, 2005.
- [52] C. Berthet, K. Raj, P. Saudan, and P. Beard, "How adeno-associated virus Rep78 protein arrests cells completely in S phase.," *Proc. Natl. Acad. Sci. U.S.A.*, vol. 102, no. 38, pp. 13634–13639, Sep. 2005.
- [53] M. Schmidt, S. Afione, and R. M. Kotin, "Adeno-associated virus type 2 Rep78 induces apoptosis through caspase activation independently of p53.," *J. Virol.*, vol. 74, no. 20, pp. 9441–9450, Oct. 2000.
- [54] N. J. Philpott, J. Gomos, K. I. Berns, and E. Falck-Pedersen, "A p5 integration efficiency element mediates Rep-dependent integration into AAVS1 at chromosome 19.," *Proc. Natl. Acad. Sci. U.S.A.*, vol. 99, no. 19, pp. 12381–12385, Sep. 2002.
- [55] B. C. Schnepf, R. L. Jensen, C.-L. Chen, P. R. Johnson, and K. R. Clark, "Characterization of adeno-associated virus genomes isolated from human tissues.," *J. Virol.*, vol. 79, no. 23, pp. 14793–14803, Dec. 2005.
- [56] R. M. Kotin, J. C. Menninger, D. C. Ward, and K. I. Berns, "Mapping and direct visualization of a region-specific viral DNA integration site on chromosome 19q13-qter.," *Genomics*, vol. 10, no. 3, pp. 831–834, Jul. 1991.
- [57] Z. Wu, A. Asokan, and R. J. Samulski, "Adeno-associated virus serotypes: vector toolkit for human gene therapy.," *Mol. Ther.*, vol. 14, no. 3, pp. 316–327, Sep. 2006.
- [58] R. M. Linden, P. Ward, C. Giraud, E. Winocour, and K. I. Berns, "Site-specific integration by adeno-associated virus.," *Proc. Natl. Acad. Sci. U.S.A.*, vol. 93, no. 21, pp. 11288–11294, Oct. 1996.
- [59] R. O. Snyder, D. S. Im, T. Ni, X. Xiao, R. J. Samulski, and N. Muzyczka, "Features of the adeno-associated virus origin involved in substrate recognition by the viral Rep protein.," *J. Virol.*, vol. 67, no. 10, pp. 6096–6104, Oct. 1993.
- [60] R. J. Samulski, X. Zhu, X. Xiao, J. D. Brook, D. E. Housman, N. Epstein, and L. A. Hunter, "Targeted integration of adeno-associated virus (AAV) into human chromosome 19.," *EMBO J.*, vol. 10, no. 12, pp. 3941–3950, Dec. 1991.
- [61] R. M. Kotin, R. M. Linden, and K. I. Berns, "Characterization of a preferred site on human chromosome 19q for integration of adeno-associated virus DNA by non-homologous recombination.," *EMBO J.*, vol. 11, no. 13, pp. 5071–5078, Dec. 1992.
- [62] M. D. Weitzman, S. R. Kyöstiö, R. M. Kotin, and R. A. Owens, "Adeno-associated virus (AAV) Rep proteins mediate complex formation between AAV DNA and its integration site in human DNA.," *Proc. Natl. Acad. Sci. U.S.A.*, vol. 91, no. 13, pp. 5808–5812, Jun. 1994.
- [63] R. van Rensburg, I. Beyer, X.-Y. Yao, H. Wang, O. Denisenko, Z.-Y. Li, D. W. Russell, D. G. Miller, P. Gregory, M. Holmes, K. Bomsztyk, and A. Lieber, "Chromatin structure of two genomic sites for targeted transgene integration in induced pluripotent stem cells and hematopoietic stem cells.," *Gene Ther.*, Mar. 2012.
- [64] D. Hüser, S. Weger, and R. Heilbronn, "Kinetics and frequency of adeno-associated virus site-specific integration into human chromosome 19 monitored by quantitative real-time PCR.," *J. Virol.*, vol. 76, no. 15, pp. 7554–7559, Aug. 2002.
- [65] D. Hüser, S. Weger, and R. Heilbronn, "Packaging of human chromosome 19-specific adeno-associated virus (AAV) integration sites in AAV virions during AAV wild-type and recombinant AAV vector production.," *J. Virol.*, vol. 77, no. 8, pp. 4881–4887, Apr. 2003.
- [66] C. C. Yang, X. Xiao, X. Zhu, D. C. Ansardi, N. D. Epstein, M. R. Frey, A. G. Matera, and R. J. Samulski, "Cellular recombination pathways and viral terminal repeat hairpin structures are sufficient for adeno-associated virus integration in vivo and in vitro.," *J. Virol.*, vol. 71, no. 12, pp. 9231–9247, Dec. 1997.

- [67] L. Pieroni, C. Fipaldini, A. Monciotti, D. Cimini, A. Sgura, E. Fattori, O. Epifano, R. Cortese, F. Palombo, and N. La Monica, "Targeted integration of adeno-associated virus-derived plasmids in transfected human cells," *Virology*, vol. 249, no. 2, pp. 249–259, Sep. 1998.
- [68] D. Pannell and J. Ellis, "Silencing of gene expression: implications for design of retrovirus vectors," *Rev. Med. Virol.*, vol. 11, no. 4, pp. 205–217, Jul. 2001.
- [69] R. D. Palmiter and R. L. Brinster, "Germ-line transformation of mice," *Annual review of genetics*, 1986.
- [70] P. Rouet, F. Smih, and M. Jasin, "Expression of a site-specific endonuclease stimulates homologous recombination in mammalian cells," *Proc. Natl. Acad. Sci. U.S.A.*, vol. 91, no. 13, pp. 6064–6068, Jun. 1994.
- [71] M. H. Porteus and D. Carroll, "Gene targeting using zinc finger nucleases," *Nat. Biotechnol.*, vol. 23, no. 8, pp. 967–973, Aug. 2005.
- [72] F. Smih, P. Rouet, P. J. Romanienko, and M. Jasin, "Double-strand breaks at the target locus stimulate gene targeting in embryonic stem cells," *Nucleic Acids Res.*, vol. 23, no. 24, pp. 5012–5019, Dec. 1995.
- [73] T. I. Cornu and T. Cathomen, "Targeted genome modifications using integrase-deficient lentiviral vectors," *Mol. Ther.*, vol. 15, no. 12, pp. 2107–2113, Dec. 2007.
- [74] Y. G. Kim, J. Cha, and S. Chandrasegaran, "Hybrid restriction enzymes: zinc finger fusions to Fok I cleavage domain," *Proc. Natl. Acad. Sci. U.S.A.*, vol. 93, no. 3, pp. 1156–1160, Feb. 1996.
- [75] F. D. Urnov, J. C. Miller, Y.-L. Lee, C. M. Beausejour, J. M. Rock, S. Augustus, A. C. Jamieson, M. H. Porteus, P. D. Gregory, and M. C. Holmes, "Highly efficient endogenous human gene correction using designed zinc-finger nucleases," *Nature*, vol. 435, no. 7042, pp. 646–651, Jun. 2005.
- [76] J.-C. Epinat, S. Arnould, P. Chames, P. Rochaix, D. Desfontaines, C. Puzin, A. Patin, A. Zanghellini, F. Pâques, and E. Lacroix, "A novel engineered meganuclease induces homologous recombination in yeast and mammalian cells," *Nucleic Acids Res.*, vol. 31, no. 11, pp. 2952–2962, Jun. 2003.
- [77] G. Silva, L. Poirrot, R. Galetto, J. Smith, G. Montoya, P. Duchateau, and F. Pâques, "Meganucleases and other tools for targeted genome engineering: perspectives and challenges for gene therapy," *Curr Gene Ther.*, vol. 11, no. 1, pp. 11–27, Feb. 2011.
- [78] A. Lombardo, P. Genovese, C. M. Beausejour, S. Colleoni, Y.-L. Lee, K. A. Kim, D. Ando, F. D. Urnov, C. Galli, P. D. Gregory, M. C. Holmes, and L. Naldini, "Gene editing in human stem cells using zinc finger nucleases and integrase-defective lentiviral vector delivery," *Nat. Biotechnol.*, vol. 25, no. 11, pp. 1298–1306, Nov. 2007.
- [79] D. Soroldoni, B. M. Hogan, and A. C. Oates, "Simple and efficient transgenesis with meganuclease constructs in zebrafish," *Methods Mol. Biol.*, vol. 546, pp. 117–130, 2009.
- [80] A. Barzel, E. Privman, M. Peeri, A. Naor, E. Shachar, D. Burstein, R. Lazary, U. Gophna, T. Pupko, and M. Kupiec, "Native homing endonucleases can target conserved genes in humans and in animal models," *Nucleic Acids Res.*, vol. 39, no. 15, pp. 6646–6659, Aug. 2011.
- [81] E. Belay, S. Dastidar, T. VandenDriessche, and M. K. L. Chuah, "Transposon-mediated gene transfer into adult and induced pluripotent stem cells," *Curr Gene Ther.*, vol. 11, no. 5, pp. 406–413, Oct. 2011.
- [82] C. A. Vink, H. B. Gaspar, R. Gabriel, M. Schmidt, R. S. McIvor, A. J. Thrasher, and W. Qasim, "Sleeping beauty transposition from nonintegrating lentivirus," *Mol. Ther.*, vol. 17, no. 7, pp. 1197–1204, Jul. 2009.
- [83] N. H. Staunstrup, B. Moldt, L. Mátés, P. Villesen, M. Jakobsen, Z. Ivics, Z. Izsvák, and J. G. Mikkelsen, "Hybrid lentivirus-transposon vectors with a random integration profile in human cells," *Mol. Ther.*, vol. 17, no. 7, pp. 1205–1214, Jul. 2009.
- [84] E. L. Aronovich, R. S. McIvor, and P. B. Hackett, "The Sleeping Beauty transposon system: a non-viral vector for gene therapy," *Hum. Mol. Genet.*, vol. 20, no. 1, pp. R14–20, Apr. 2011.
- [85] A. Akopian and W. Marshall Stark, "Site-specific DNA recombinases as instruments for genomic surgery," *Adv. Genet.*, vol. 55, pp. 1–23, 2005.
- [86] J. R. Heil, J. Cheng, and T. C. Charles, "Site-specific bacterial chromosome engineering:  $\Phi$ C31 integrase mediated cassette exchange (IMCE)," *J Vis Exp*, no. 61, 2012.

- [87] H. M. Thorpe and M. C. Smith, "In vitro site-specific integration of bacteriophage DNA catalyzed by a recombinase of the resolvase/invertase family.," *Proc. Natl. Acad. Sci. U.S.A.*, vol. 95, no. 10, pp. 5505–5510, May 1998.
- [88] K. Abremski and R. Hoess, "Bacteriophage P1 site-specific recombination. Purification and properties of the Cre recombinase protein.," *J. Biol. Chem.*, vol. 259, no. 3, pp. 1509–1514, Feb. 1984.
- [89] B. J. Andrews, G. A. Proteau, L. G. Beatty, and P. D. Sadowski, "The FLP recombinase of the 2 micron circle DNA of yeast: interaction with its target sequences.," *Cell*, vol. 40, no. 4, pp. 795–803, Apr. 1985.
- [90] C. S. Branda and S. M. Dymecki, "Talking about a revolution: The impact of site-specific recombinases on genetic analyses in mice.," *Dev. Cell*, vol. 6, no. 1, pp. 7–28, Jan. 2004.
- [91] S. Turan and J. Bode, "Site-specific recombinases: from tag-and-target- to tag-and-exchange-based genomic modifications.," *FASEB J.*, vol. 25, no. 12, pp. 4088–4107, Dec. 2011.
- [92] A. Baer and J. Bode, "Coping with kinetic and thermodynamic barriers: RMCE, an efficient strategy for the targeted integration of transgenes.," *Curr. Opin. Biotechnol.*, 2001.
- [93] H. Albert, E. C. Dale, E. Lee, and D. W. Ow, "Site-specific integration of DNA into wild-type and mutant lox sites placed in the plant genome.," *Plant J.*, vol. 7, no. 4, pp. 649–659, Apr. 1995.
- [94] G. Lee and I. Saito, "Role of nucleotide sequences of loxP spacer region in Cre-mediated recombination.," *Gene*, vol. 216, no. 1, pp. 55–65, Aug. 1998.
- [95] A. F. Kolb, "Selection-marker-free modification of the murine beta-casein gene using a lox2272 [correction of lox2722] site.," *Anal. Biochem.*, vol. 290, no. 2, pp. 260–271, Mar. 2001.
- [96] R. M. Blaese, K. W. Culver, A. D. Miller, C. S. Carter, T. Fleisher, M. Clerici, G. Shearer, L. Chang, Y. Chiang, P. Tolstoshev, J. J. Greenblatt, S. A. Rosenberg, H. Klein, M. Berger, C. A. Mullen, W. J. Ramsey, L. Muul, R. A. Morgan, and W. F. Anderson, "T lymphocyte-directed gene therapy for ADA-SCID: initial trial results after 4 years.," *Science*, vol. 270, no. 5235, pp. 475–480, Oct. 1995.
- [97] M. K. Brenner, D. R. Rill, R. C. Moen, R. A. Krance, J. Mirro, W. F. Anderson, and J. N. Ihle, "Gene-marking to trace origin of relapse after autologous bone-marrow transplantation.," *Lancet*, vol. 341, no. 8837, pp. 85–86, Jan. 1993.
- [98] J. A. Nolte, E. M. Smogorzewska, and D. B. Kohn, "Analysis of optimal conditions for retroviral-mediated transduction of primitive human hematopoietic cells.," *Blood*, 1995.
- [99] S. Hacein-Bey-Abina, C. Von Kalle, M. Schmidt, M. P. McCormack, N. Wulffraat, P. Leboulch, A. Lim, C. S. Osborne, R. Pawliuk, E. Morillon, R. Sorensen, A. Forster, P. Fraser, J. I. Cohen, G. de Saint Basile, I. Alexander, U. Wintergerst, T. Frebourg, A. Aurias, D. Stoppa-Lyonnet, S. Romana, I. Radford-Weiss, F. Gross, F. Valensi, E. Delabesse, E. Macintyre, F. Sigaux, J. Soulier, L. E. Leiva, M. Wissler, C. Prinz, T. H. Rabbitts, F. Le Deist, A. Fischer, and M. Cavazzana-Calvo, "LMO2-associated clonal T cell proliferation in two patients after gene therapy for SCID-X1.," *Science*, vol. 302, no. 5644, pp. 415–419, Oct. 2003.
- [100] C. Baum, O. Kustikova, U. Modlich, and Z. Li, "Mutagenesis and oncogenesis by chromosomal insertion of gene transfer vectors.," *Human gene ...*, 2006.
- [101] L. Naldini, U. Blömer, P. Gallay, D. Ory, R. Mulligan, F. H. Gage, I. M. Verma, and D. Trono, "In vivo gene delivery and stable transduction of nondividing cells by a lentiviral vector.," *Science*, vol. 272, no. 5259, pp. 263–267, Apr. 1996.
- [102] R. Zufferey, T. Dull, R. J. Mandel, A. Bukovsky, D. Quiroz, L. Naldini, and D. Trono, "Self-inactivating lentivirus vector for safe and efficient in vivo gene delivery.," *J. Virol.*, vol. 72, no. 12, pp. 9873–9880, Dec. 1998.
- [103] Y. Wu, "HIV-1 gene expression: lessons from provirus and non-integrated DNA.," *Retrovirology*, vol. 1, p. 13, 2004.
- [104] A. D. Frankel and J. Young, "HIV-1: fifteen proteins and an RNA.," *Annual review of biochemistry*, 1998.
- [105] J. A. G. Briggs, M. N. Simon, I. Gross, H.-G. Kräusslich, S. D. Fuller, V. M. Vogt, and M. C. Johnson, "The stoichiometry of Gag protein in HIV-1.," *Nat. Struct. Mol. Biol.*, vol. 11, no. 7, pp. 672–675, Jul. 2004.
- [106] T. Sakuma, M. A. Barry, and Y. Ikeda, "Lentiviral vectors: basic to translational.," *Biochem. J.*, vol. 443, no. 3, pp. 603–618, May 2012.

- [107] C. J. Buchholz, M. D. Mühlebach, and K. Cichutek, "Lentiviral vectors with measles virus glycoproteins - dream team for gene transfer?," *Trends in Biotechnology*, vol. 27, no. 5, pp. 259–265, May 2009.
- [108] J. C. Burns, T. Friedmann, W. Driever, M. Burrascano, and J. K. Yee, "Vesicular stomatitis virus G glycoprotein pseudotyped retroviral vectors: concentration to very high titer and efficient gene transfer into mammalian and nonmammalian cells.," *Proc. Natl. Acad. Sci. U.S.A.*, vol. 90, no. 17, pp. 8033–8037, Sep. 1993.
- [109] R. Mann, R. C. Mulligan, and D. Baltimore, "Construction of a retrovirus packaging mutant and its use to produce helper-free defective retrovirus.," *Cell*, 1983.
- [110] S. Watanabe and H. M. Temin, "Construction of a helper cell line for avian reticuloendotheliosis virus cloning vectors.," *Mol. Cell. Biol.*, vol. 3, no. 12, pp. 2241–2249, Dec. 1983.
- [111] A. D. Miller and G. J. Rosman, "Improved retroviral vectors for gene transfer and expression," *BioTechniques*, 1989.
- [112] K.-I. Joo and P. Wang, "Visualization of targeted transduction by engineered lentiviral vectors.," *Gene Ther.*, vol. 15, no. 20, pp. 1384–1396, Oct. 2008.
- [113] E. Vigna and L. Naldini, "Lentiviral vectors: excellent tools for experimental gene transfer and promising candidates for gene therapy," *J Gene Med*, 2000.
- [114] D. Trono, "Lentiviral vectors: turning a deadly foe into a therapeutic agent," *Gene Ther.*, 2000.
- [115] Y. Suzuki and R. Craigie, "The road to chromatin - nuclear entry of retroviruses.," *Nat. Rev. Microbiol.*, vol. 5, no. 3, pp. 187–196, Mar. 2007.
- [116] R. Craigie, "HIV integrase, a brief overview from chemistry to therapeutics," *J. Biol. Chem.*, 2001.
- [117] S. Pang, Y. Koyanagi, S. Miles, C. Wiley, H. V. Vinters, and I. S. Chen, "High levels of unintegrated HIV-1 DNA in brain tissue of AIDS dementia patients.," *Nature*, vol. 343, no. 6253, pp. 85–89, Jan. 1990.
- [118] Y. Wu, "The second chance story of HIV-1 DNA: Unintegrated? Not a problem!," *Retrovirology*, vol. 5, p. 61, 2008.
- [119] C. M. Farnet and W. A. Haseltine, "Circularization of human immunodeficiency virus type 1 DNA in vitro.," *J. Virol.*, vol. 65, no. 12, pp. 6942–6952, Dec. 1991.
- [120] L. Li, J. M. Olvera, K. E. Yoder, R. S. Mitchell, S. L. Butler, M. Lieber, S. L. Martin, and F. D. Bushman, "Role of the non-homologous DNA end joining pathway in the early steps of retroviral infection.," *EMBO J.*, vol. 20, no. 12, pp. 3272–3281, Jun. 2001.
- [121] A. Engelman, K. Mizuuchi, and R. Craigie, "HIV-1 DNA integration: mechanism of viral DNA cleavage and DNA strand transfer," *Cell*, 1991.
- [122] S. L. Butler, M. S. Hansen, and F. D. Bushman, "A quantitative assay for HIV DNA integration in vivo.," *Nat. Med.*, vol. 7, no. 5, pp. 631–634, May 2001.
- [123] V. Zennou, C. Petit, D. Guetard, U. Nerhbass, L. Montagnier, and P. Charneau, "HIV-1 genome nuclear import is mediated by a central DNA flap.," *Cell*, vol. 101, no. 2, pp. 173–185, Apr. 2000.
- [124] I. Raineri and H. P. Senn, "HIV-1 promotor insertion revealed by selective detection of chimeric provirus-host gene transcripts.," *Nucleic Acids Res.*, vol. 20, no. 23, pp. 6261–6266, Dec. 1992.
- [125] B. Shiramizu, B. G. Herndier, and M. S. McGrath, "Identification of a common clonal human immunodeficiency virus integration site in human immunodeficiency virus-associated lymphomas.," *Cancer Res.*, vol. 54, no. 8, pp. 2069–2072, Apr. 1994.
- [126] S. J. Nightingale, R. P. Hollis, K. A. Pepper, D. Petersen, X.-J. Yu, C. Yang, I. Bahner, and D. B. Kohn, "Transient gene expression by nonintegrating lentiviral vectors.," *Mol. Ther.*, vol. 13, no. 6, pp. 1121–1132, Jun. 2006.
- [127] S. Philippe, C. Sarkis, M. Barkats, H. Mammeri, C. Ladroue, C. Petit, J. Mallet, and C. Serguera, "Lentiviral vectors with a defective integrase allow efficient and sustained transgene expression in vitro and in vivo.," *Proc. Natl. Acad. Sci. U.S.A.*, vol. 103, no. 47, pp. 17684–17689, Nov. 2006.
- [128] R. J. Yáñez-Muñoz, K. S. Balaggan, A. MacNeil, S. J. Howe, M. Schmidt, A. J. Smith, P. Buch, R. E. MacLaren, P. N. Anderson, S. E. Barker, Y. Duran, C. Bartholomae, C. von Kalle, J. R. Heckenlively, C. Kinnon, R. R. Ali, and A. J. Thrasher, "Effective gene therapy with nonintegrating lentiviral vectors.," *Nat. Med.*, vol. 12, no. 3, pp. 348–353, Mar. 2006.
- [129] P. Gallay, T. Hope, D. Chin, and D. Trono, "HIV-1 infection of nondividing cells through the recognition of integrase by the importin/karyopherin pathway.," *Proc. Natl. Acad. Sci. U.S.A.*, vol. 94, no. 18, pp. 9825–9830, Sep. 1997.

- [130] A. Engelman, "In Vivo Analysis of Retroviral Integrase Structure and Function," *Advances in virus research*, 1999.
- [131] A. Engelman, G. Englund, J. M. Orenstein, M. A. Martin, and R. Craigie, "Multiple effects of mutations in human immunodeficiency virus type 1 integrase on viral replication," *J. Virol.*, vol. 69, no. 5, pp. 2729–2736, May 1995.
- [132] N. J. Philpott and A. J. Thrasher, "Use of nonintegrating lentiviral vectors for gene therapy," *Hum. Gene Ther.*, 2007.
- [133] A. Engelman, F. D. Bushman, and R. Craigie, "Identification of discrete functional domains of HIV-1 integrase and their organization within an active multimeric complex," *EMBO J.*, vol. 12, no. 8, pp. 3269–3275, Aug. 1993.
- [134] A. Engelman and R. Craigie, "Identification of conserved amino acid residues critical for human immunodeficiency virus type 1 integrase function in vitro," *J. Virol.*, vol. 66, no. 11, pp. 6361–6369, Nov. 1992.
- [135] M. Nowotny, "Retroviral integrase superfamily: the structural perspective," *EMBO Rep.*, vol. 10, no. 2, pp. 144–151, Feb. 2009.
- [136] M. B. Banasik and P. B. McCray, "Integrase-defective lentiviral vectors: progress and applications," *Gene Ther.*, vol. 17, no. 2, pp. 150–157, Feb. 2010.
- [137] T. C. Pierson, T. L. Kieffer, C. T. Ruff, C. Buck, S. J. Gange, and R. F. Siliciano, "Intrinsic stability of episomal circles formed during human immunodeficiency virus type 1 replication," *J. Virol.*, vol. 76, no. 8, pp. 4138–4144, Apr. 2002.
- [138] J. J. Harrington, G. Van Bokkelen, R. W. Mays, K. Gustashaw, and H. F. Willard, "Formation of de novo centromeres and construction of first-generation human artificial microchromosomes," *Nature Genetics*, vol. 15, no. 4, pp. 345–355, Apr. 1997.
- [139] M. M. P. Lufino, P. A. H. Edser, and R. Wade-Martins, "Advances in high-capacity extrachromosomal vector technology: episomal maintenance, vector delivery, and transgene expression," *Mol. Ther.*, vol. 16, no. 9, pp. 1525–1538, Sep. 2008.
- [140] D. J. Glover, H. J. Lipps, and D. A. Jans, "Towards safe, non-viral therapeutic gene expression in humans," *Nat Rev Genet*, 2005.
- [141] E. Lin and J. Nemunaitis, "Oncolytic viral therapies," *Cancer Gene Ther.*, vol. 11, no. 10, pp. 643–664, Oct. 2004.
- [142] J. L. Yates, N. Warren, and B. Sugden, "Stable replication of plasmids derived from Epstein–Barr virus in various mammalian cells," 1985.
- [143] J. Vargas Jr, G. L. Gusella, and V. Najfeld, "Novel integrase-defective lentiviral episomal vectors for gene transfer," *Human gene ...*, 2004.
- [144] D. M. Gilbert, "Making sense of eukaryotic DNA replication origins," *Science*, 2001.
- [145] Q. Li, K. R. Peterson, X. Fang, and G. Stamatoyannopoulos, "Locus control regions," *Blood*, vol. 100, no. 9, pp. 3077–3086, Nov. 2002.
- [146] G. B. Price, M. Allarakhia, N. Cossons, T. Nielsen, M. Diaz-Perez, P. Friedlander, L. Tao, and M. Zannis-Hadjopoulos, "Identification of a cis-element that determines autonomous DNA replication in eukaryotic cells," *J. Biol. Chem.*, vol. 278, no. 22, pp. 19649–19659, May 2003.
- [147] S. M. Stoll and M. P. Calos, "Extrachromosomal plasmid vectors for gene therapy," *Curr. Opin. Mol. Ther.*, vol. 4, no. 4, pp. 299–305, Aug. 2002.
- [148] C. Piechaczek, C. Fetzter, A. Baiker, J. Bode, and H. J. Lipps, "A vector based on the SV40 origin of replication and chromosomal S/MARs replicates episomally in CHO cells," *Nucleic Acids Res.*, vol. 27, no. 2, pp. 426–428, Jan. 1999.
- [149] A. Baiker, C. Maercker, C. Piechaczek, S. B. Schmidt, J. Bode, C. Benham, and H. J. Lipps, "Mitotic stability of an episomal vector containing a human scaffold/matrix-attached region is provided by association with nuclear matrix," *Nat. Cell Biol.*, vol. 2, no. 3, pp. 182–184, Mar. 2000.
- [150] I. M. Stehle, M. F. Scinteie, A. Baiker, A. C. W. Jenke, and H. J. Lipps, "Exploiting a minimal system to study the epigenetic control of DNA replication: the interplay between transcription and replication," *Chromosome Res.*, vol. 11, no. 5, pp. 413–421, 2003.
- [151] A. M. Darquet, R. Rangara, P. Kreiss, B. Schwartz, S. Naimi, P. Delaère, J. Crouzet, and D. Scherman, "Minicircle: an improved DNA molecule for in vitro and in vivo gene transfer," *Gene Ther.*, vol. 6, no. 2, pp. 209–218, Feb. 1999.

- [152] R. Darteil, M. Wang, M. Latta-Mahieu, A. Caron, A. Mahfoudi, B. Staels, and V. Thuillier, "Efficient gene regulation by PPAR gamma and thiazolidinediones in skeletal muscle and heart," *Mol. Ther.*, vol. 6, no. 2, pp. 265–271, Aug. 2002.
- [153] L. T. Feldman, "Problems in the use of herpes simplex virus as a vector," *Int. Rev. Neurobiol.*, vol. 55, pp. 99–109, 2003.
- [154] N. Tatsis and H. C. J. Ertl, "Adenoviruses as vaccine vectors," *Mol. Ther.*, vol. 10, no. 4, pp. 616–629, Oct. 2004.
- [155] H. Hoshiya, Y. Kazuki, S. Abe, M. Takiguchi, N. Kajitani, Y. Watanabe, T. Yoshino, Y. Shirayoshi, K. Higaki, G. Messina, G. Cossu, and M. Oshimura, "A highly stable and nonintegrated human artificial chromosome (HAC) containing the 2.4 Mb entire human dystrophin gene," *Mol. Ther.*, vol. 17, no. 2, pp. 309–317, Feb. 2009.
- [156] R. Lu, N. Nakajima, W. Hofmann, M. Benkirane, K.-T. Jeang, J. Sodroski, A. Engelman, and K. Teh-Jeang, "Simian virus 40-based replication of catalytically inactive human immunodeficiency virus type 1 integrase mutants in nonpermissive T cells and monocyte-derived macrophages," *J. Virol.*, vol. 78, no. 2, pp. 658–668, Jan. 2004.
- [157] R. Langer, "Drug delivery and targeting," *Nature*, vol. 392, no. 6679, pp. 5–10, Apr. 1998.
- [158] J. S. Wadia and S. F. Dowdy, "Protein transduction technology," *Curr. Opin. Biotechnol.*, vol. 13, no. 1, pp. 52–56, Feb. 2002.
- [159] A. D. Frankel and C. O. Pabo, "Cellular uptake of the tat protein from human immunodeficiency virus," *Cell*, vol. 55, no. 6, pp. 1189–1193, Dec. 1988.
- [160] D. Derossi, A. H. Joliot, G. Chassaing, and A. Prochiantz, "The third helix of the Antennapedia homeodomain translocates through biological membranes," *J. Biol. Chem.*, vol. 269, no. 14, pp. 10444–10450, Apr. 1994.
- [161] S. R. Schwarze, K. A. Hruska, and S. F. Dowdy, "Protein transduction: unrestricted delivery into all cells?," *Trends Cell Biol.*, vol. 10, no. 7, pp. 290–295, Jul. 2000.
- [162] X. Wu, H. Liu, H. Xiao, J. Kim, P. Seshiah, G. Natsoulis, J. D. Boeke, B. H. Hahn, and J. C. Kappes, "Targeting foreign proteins to human immunodeficiency virus particles via fusion with Vpr and Vpx," *J. Virol.*, vol. 69, no. 6, pp. 3389–3398, Jun. 1995.
- [163] N. Link, C. Aubel, J. M. Kelm, R. R. Marty, D. Greber, V. Djonov, J. Bourhis, W. Weber, and M. Fussenegger, "Therapeutic protein transduction of mammalian cells and mice by nucleic acid-free lentiviral nanoparticles," *Nucleic Acids Res.*, vol. 34, no. 2, p. e16, 2006.
- [164] W. Paxton, R. I. Connor, and N. R. Landau, "Incorporation of Vpr into human immunodeficiency virus type 1 virions: requirement for the p6 region of gag and mutational analysis," *J. Virol.*, vol. 67, no. 12, pp. 7229–7237, Dec. 1993.
- [165] K. Muthumani, S. Kudchodkar, E. Papasavvas, L. J. Montaner, D. B. Weiner, and V. Ayyavoo, "HIV-1 Vpr regulates expression of beta chemokines in human primary lymphocytes and macrophages," *J. Leukoc. Biol.*, vol. 68, no. 3, pp. 366–372, Sep. 2000.
- [166] H. S. Moon and J.-S. Yang, "Role of HIV Vpr as a regulator of apoptosis and an effector on bystander cells," *Mol. Cells*, vol. 21, no. 1, pp. 7–20, Feb. 2006.
- [167] R. Zufferey, D. Nagy, R. J. Mandel, L. Naldini, and D. Trono, "Multiply attenuated lentiviral vector achieves efficient gene delivery in vivo," *Nat. Biotechnol.*, vol. 15, no. 9, pp. 871–875, Sep. 1997.
- [168] T. Jacks, M. D. Power, F. R. Masiarz, P. A. Luciw, P. J. Barr, and H. E. Varmus, "Characterization of ribosomal frameshifting in HIV-1 gag-pol expression," *Nature*, vol. 331, no. 6153, pp. 280–283, Jan. 1988.
- [169] W. Hübner, P. Chen, A. Del Portillo, Y. Liu, R. E. Gordon, and B. K. Chen, "Sequence of human immunodeficiency virus type 1 (HIV-1) Gag localization and oligomerization monitored with live confocal imaging of a replication-competent, fluorescently tagged HIV-1," *J. Virol.*, vol. 81, no. 22, pp. 12596–12607, Nov. 2007.
- [170] A. D. Adamson, D. Jackson, and J. R. E. Davis, "Novel approaches to in vitro transgenesis," *J. Endocrinol.*, vol. 208, no. 3, pp. 193–206, Mar. 2011.
- [171] A. Presser, M. B. Elowitz, M. Kellis, and R. Kishony, "The evolutionary dynamics of the *Saccharomyces cerevisiae* protein interaction network after duplication," *Proceedings of the National Academy of Sciences*, vol. 105, no. 3, pp. 950–954, Jan. 2008.



- [172] J. F. Sabl and S. Henikoff, "Copy number and orientation determine the susceptibility of a gene to silencing by nearby heterochromatin in *Drosophila*," *Genetics*, vol. 142, no. 2, pp. 447–458, Feb. 1996.
- [173] Y. Zhang and B. Oliver, "Dosage compensation goes global," *Curr. Opin. Genet. Dev.*, vol. 17, no. 2, pp. 113–120, Apr. 2007.
- [174] Y. Mileyko, "Bifurcation analysis of gene regulatory circuits subject to copy number variation," 2010.
- [175] N. Dillon and R. Festenstein, "Unravelling heterochromatin: competition between positive and negative factors regulates accessibility," *Trends Genet.*, vol. 18, no. 5, pp. 252–258, May 2002.
- [176] V. Mutskov and G. Felsenfeld, "Silencing of transgene transcription precedes methylation of promoter DNA and histone H3 lysine 9," *EMBO J.*, vol. 23, no. 1, pp. 138–149, Jan. 2004.
- [177] A. C. W. Jenke, I. M. Stehle, F. Herrmann, T. Eisenberger, A. Baiker, J. Bode, F. O. Fackelmayer, and H. J. Lipps, "Nuclear scaffold/matrix attached region modules linked to a transcription unit are sufficient for replication and maintenance of a mammalian episome," *Proc. Natl. Acad. Sci. U.S.A.*, vol. 101, no. 31, pp. 11322–11327, Aug. 2004.
- [178] L. J. Burke and A. Baniahmad, "Co-repressors 2000," *The FASEB Journal*, 2000.
- [179] M. O. Clements, A. Godfrey, J. Crossley, S. J. Wilson, Y. Takeuchi, and C. Boshoff, "Lentiviral manipulation of gene expression in human adult and embryonic stem cells," *Tissue Eng.*, vol. 12, no. 7, pp. 1741–1751, Jul. 2006.
- [180] D. Wirth, L. Gama-Norton, P. Riemer, and U. Sandhu, "Road to precision: recombinase-based targeting technologies for genome engineering," *Current opinion in ...*, 2007.
- [181] R. Tomanin and M. Scarpa, "Why do we need new gene therapy viral vectors? Characteristics, limitations and future perspectives of viral vector transduction," *Curr Gene Ther.*, vol. 4, no. 4, pp. 357–372, Dec. 2004.
- [182] N. Mütter, N. Noske, and A. Ehrhardt, "Viral hybrid vectors for somatic integration - are they the better solution?," *Viruses*, vol. 1, no. 3, pp. 1295–1324, Dec. 2009.
- [183] R. R. Ahangarani, W. Janssens, V. Carlier, L. Vanderelst, T. VandenDriessche, M. Chuah, M. Jacquemin, and J.-M. Saint-Remy, "Retroviral vectors induce epigenetic chromatin modifications and IL-10 production in transduced B cells via activation of toll-like receptor 2," *Mol. Ther.*, vol. 19, no. 4, pp. 711–722, Apr. 2011.
- [184] F. D. Bushman, "Tethering human immunodeficiency virus 1 integrase to a DNA site directs integration to nearby sequences," *Proc. Natl. Acad. Sci. U.S.A.*, vol. 91, no. 20, pp. 9233–9237, Sep. 1994.
- [185] R. A. Katz, G. Merkel, and A. M. Skalka, "Targeting of retroviral integrase by fusion to a heterologous DNA binding domain: in vitro activities and incorporation of a fusion protein into viral particles," *Virology*, vol. 217, no. 1, pp. 178–190, Mar. 1996.
- [186] W. Tan, Z. Dong, T. A. Wilkinson, C. F. Barbas, and S. A. Chow, "Human immunodeficiency virus type 1 incorporated with fusion proteins consisting of integrase and the designed polydactyl zinc finger protein E2C can bias integration of viral DNA into a predetermined chromosomal region in human cells," *J. Virol.*, vol. 80, no. 4, pp. 1939–1948, Feb. 2006.
- [187] B. Moldt, N. H. Staunstrup, M. Jakobsen, R. J. Yáñez-Muñoz, and J. G. Mikkelsen, "Genomic insertion of lentiviral DNA circles directed by the yeast Flp recombinase," *BMC Biotechnol.*, vol. 8, p. 60, 2008.
- [188] R. Torres, A. García, M. Payá, and J. C. Ramirez, "Non-integrative lentivirus drives high-frequency cre-mediated cassette exchange in human cells," *PLoS ONE*, vol. 6, no. 5, p. e19794, 2011.
- [189] J. G. Mikkelsen, S. R. Yant, L. Meuse, Z. Huang, H. Xu, and M. A. Kay, "Helper-Independent Sleeping Beauty transposon-transposase vectors for efficient nonviral gene delivery and persistent gene expression in vivo," *Mol. Ther.*, vol. 8, no. 4, pp. 654–665, Oct. 2003.
- [190] V. Pattanayak, C. L. Ramirez, J. K. Joung, and D. R. Liu, "Revealing off-target cleavage specificities of zinc-finger nucleases by in vitro selection," *Nat. Methods*, vol. 8, no. 9, pp. 765–770, Sep. 2011.
- [191] S. Turan, M. Galla, E. Ernst, J. Qiao, C. Voelkel, B. Schiedlmeier, C. Zehe, and J. Bode, "Recombinase-mediated cassette exchange (RMCE): traditional concepts and current challenges," *J. Mol. Biol.*, vol. 407, no. 2, pp. 193–221, Mar. 2011.

- [192] J. Y. Qin, L. Zhang, K. L. Clift, I. Hultur, A. P. Xiang, B.-Z. Ren, and B. T. Lahn, "Systematic comparison of constitutive promoters and the doxycycline-inducible promoter.," *PLoS ONE*, vol. 5, no. 5, p. e10611, 2010.
- [193] B. N. Fields, D. M. Knipe, and P. M. Howley, *Fields' virology*. Lippincott Williams & Wilkins, 2007.
- [194] B. Fehse, O. S. Kustikova, M. Bubenheim, and C. Baum, "Pois(s)on--it's a question of dose...," *Gene Ther*, vol. 11, no. 11, pp. 879–881, Jun. 2004.
- [195] S. Irion, H. Luche, P. Gadue, H. J. Fehling, M. Kennedy, and G. Keller, "Identification and targeting of the ROSA26 locus in human embryonic stem cells.," *Nat. Biotechnol.*, vol. 25, no. 12, pp. 1477–1482, Dec. 2007.
- [196] R. C. DeKolver, V. M. Choi, E. A. Moehle, D. E. Paschon, D. Hockemeyer, S. H. Meijssing, Y. Sancak, X. Cui, E. J. Steine, J. C. Miller, P. Tam, V. V. Bartsevich, X. Meng, I. Rupniewski, S. M. Gopalan, H. C. Sun, K. J. Pitz, J. M. Rock, L. Zhang, G. D. Davis, E. J. Rebar, I. M. Cheeseman, K. R. Yamamoto, D. M. Sabatini, R. Jaenisch, P. D. Gregory, and F. D. Urnov, "Functional genomics, proteomics, and regulatory DNA analysis in isogenic settings using zinc finger nuclease-driven transgenesis into a safe harbor locus in the human genome.," *Genome Res.*, vol. 20, no. 8, pp. 1133–1142, Aug. 2010.
- [197] D. Hockemeyer, F. Soldner, C. Beard, Q. Gao, M. Mitalipova, R. C. DeKolver, G. E. Katibah, R. Amora, E. A. Boydston, B. Zeitler, X. Meng, J. C. Miller, L. Zhang, E. J. Rebar, P. D. Gregory, F. D. Urnov, and R. Jaenisch, "Efficient targeting of expressed and silent genes in human ESCs and iPSCs using zinc-finger nucleases.," *Nat. Biotechnol.*, vol. 27, no. 9, pp. 851–857, Sep. 2009.
- [198] K. Wanisch and R. J. Yáñez-Muñoz, "Integration-deficient lentiviral vectors: a slow coming of age.," *Mol. Ther.*, vol. 17, no. 8, pp. 1316–1332, Aug. 2009.
- [199] K. Burger, K. Matt, N. Kieser, D. Gebhard, and J. Bergemann, "A modified fluorimetric host cell reactivation assay to determine the repair capacity of primary keratinocytes, melanocytes and fibroblasts.," *BMC Biotechnol.*, vol. 10, p. 46, 2010.
- [200] J. N. Kabalin, D. M. Peehl, and T. A. Stamey, "Clonal growth of human prostatic epithelial cells is stimulated by fibroblasts.," *Prostate*, vol. 14, no. 3, pp. 251–263, 1989.
- [201] J. M. McAllister and P. J. Hornsby, "Improved clonal and nonclonal growth of human, rat and bovine adrenocortical cells in culture.," *In Vitro Cell. Dev. Biol.*, vol. 23, no. 10, pp. 677–685, Oct. 1987.
- [202] T. Ohno and M. Yamada, "Growth stimulation by direct cell-cell contact (contact stimulation) in normal human lung fibroblasts.," *Biomed. Pharmacother.*, vol. 38, no. 7, pp. 337–343, 1984.
- [203] T. Dull, R. Zufferey, M. Kelly, and R. J. Mandel, "A third-generation lentivirus vector with a conditional packaging system.," *Journal of ...*, 1998.
- [204] C. Voelkel, M. Galla, T. Maetzig, E. Warlich, J. Kuehle, D. Zychlinski, J. Bode, T. Cantz, A. Schambach, and C. Baum, "Protein transduction from retroviral Gag precursors.," *Proceedings of the National Academy of Sciences*, vol. 107, no. 17, pp. 7805–7810, Apr. 2010.
- [205] B. Müller, J. Daecke, O. T. Fackler, M. T. Dittmar, H. Zentgraf, and H.-G. Kräusslich, "Construction and characterization of a fluorescently labeled infectious human immunodeficiency virus type 1 derivative.," *J. Virol.*, vol. 78, no. 19, pp. 10803–10813, Oct. 2004.
- [206] M. D. Muzumdar, B. Tasic, K. Miyamichi, L. Li, and L. Luo, "A global double-fluorescent Cre reporter mouse.," *Genesis*, vol. 45, no. 9, pp. 593–605, Sep. 2007.
- [207] H. Hamilton, J. Gomos, K. I. Berns, and E. Falck-Pedersen, "Adeno-associated virus site-specific integration and AAVS1 disruption.," *J. Virol.*, vol. 78, no. 15, pp. 7874–7882, Aug. 2004.
- [208] K. Nehlsen, S. Broll, and J. Bode, "Replicating minicircles: Generation of nonviral episomes for the efficient modification of dividing cells," *Gene Ther Mol Biol*, 2006.
- [209] D. Kitsberg, S. Selig, I. Keshet, and H. Cedar, "Replication structure of the human beta-globin gene domain.," *Nature*, vol. 366, no. 6455, pp. 588–590, Dec. 1993.
- [210] C. McWhinney and M. Leffak, "Autonomous replication of a DNA fragment containing the chromosomal replication origin of the human c-myc gene.," *Nucleic Acids Res.*, vol. 18, no. 5, pp. 1233–1242, Mar. 1990.
- [211] F. D. Araujo, J. D. Knox, S. Ramchandani, R. Pelletier, P. Bigey, G. Price, M. Szyf, and M. Zannis-Hadjopoulos, "Identification of initiation sites for DNA replication in the human dnmt1 (DNA-methyltransferase) locus.," *J. Biol. Chem.*, vol. 274, no. 14, pp. 9335–9341, Apr. 1999.
- [212] J. Bode, C. P. Fetzter, K. Nehlsen, and M. Scinteie, "The hitchhiking principle: optimizing episomal vectors for the use in gene therapy and biotechnology.," *Gene Ther Mol ...*, 2001.

- [213] D. Klehr, K. Maass, and J. Bode, "Scaffold-attached regions from the human interferon beta domain can be used to enhance the stable expression of genes under the control of various promoters.," *Biochemistry*, vol. 30, no. 5, pp. 1264–1270, Feb. 1991.
- [214] J. Bode, C. Benham, A. Knopp, and C. Mielke, "Transcriptional augmentation: modulation of gene expression by scaffold/matrix-attached regions (S/MAR elements).," *Crit. Rev. Eukaryot. Gene Expr.*, vol. 10, no. 1, pp. 73–90, 2000.
- [215] Q. Dang, J. Auten, and I. Plavec, "Human beta interferon scaffold attachment region inhibits de novo methylation and confers long-term, copy number-dependent expression to a retroviral vector.," *J. Virol.*, vol. 74, no. 6, pp. 2671–2678, Mar. 2000.
- [216] J. Bode, Y. Kohwi, L. Dickinson, T. Joh, D. Klehr, and C. Mielke, "Biological significance of unwinding capability of nuclear matrix-associating DNAs," *Science*, 1992.
- [217] A. Ramezani, T. S. Hawley, and R. G. Hawley, "Performance- and safety-enhanced lentiviral vectors containing the human interferon-beta scaffold attachment region and the chicken beta-globin insulator.," *Blood*, vol. 101, no. 12, pp. 4717–4724, Jun. 2003.
- [218] L. D. Mesner, J. L. Hamlin, and P. A. Dijkwel, "The matrix attachment region in the Chinese hamster dihydrofolate reductase origin of replication may be required for local chromatid separation.," *Proc. Natl. Acad. Sci. U.S.A.*, vol. 100, no. 6, pp. 3281–3286, Mar. 2003.
- [219] P. N. Cockerill and W. T. Garrard, "Chromosomal loop anchorage of the kappa immunoglobulin gene occurs next to the enhancer in a region containing topoisomerase II sites.," *Cell*, vol. 44, no. 2, pp. 273–282, Jan. 1986.
- [220] S. Stein, M. G. Ott, S. Schultze-Strasser, A. Jauch, B. Burwinkel, A. Kinner, M. Schmidt, A. Krämer, J. Schwäble, H. Glimm, U. Koehl, C. Preiss, C. Ball, H. Martin, G. Göhring, K. Schwarzwaelder, W.-K. Hofmann, K. Karakaya, S. Tchatchou, R. Yang, P. Reinecke, K. Kühlcke, B. Schlegelberger, A. J. Thrasher, D. Hoelzer, R. Seger, C. von Kalle, and M. Grez, "Genomic instability and myelodysplasia with monosomy 7 consequent to EVI1 activation after gene therapy for chronic granulomatous disease.," *Nat. Med.*, vol. 16, no. 2, pp. 198–204, Feb. 2010.
- [221] N. M. Kane, A. Nowrouzi, S. Mukherjee, M. P. Blundell, J. A. Greig, W. K. Lee, M. D. Houslay, G. Milligan, J. C. Mountford, C. von Kalle, M. Schmidt, A. J. Thrasher, and A. H. Baker, "Lentivirus-mediated reprogramming of somatic cells in the absence of transgenic transcription factors.," *Mol. Ther.*, vol. 18, no. 12, pp. 2139–2145, Dec. 2010.
- [222] T. Schlake and J. Bode, "Use of mutated FLP recognition target (FRT) sites for the exchange of expression cassettes at defined chromosomal loci," *Biochemistry*, 1994.
- [223] B. Bethke and B. Sauer, "Segmental genomic replacement by Cre-mediated recombination: genotoxic stress activation of the p53 promoter in single-copy transformants," *Nucleic Acids Res.*, 1997.
- [224] J. Seibler, D. Schübeler, S. Fiering, and M. Groudine, "DNA cassette exchange in ES cells mediated by Flp recombinase: an efficient strategy for repeated modification of tagged loci by marker-free constructs," *Biochemistry*, 1998.
- [225] A. Oumard, J. Qiao, T. Jostock, J. Li, and J. Bode, "Recommended method for chromosome exploitation: RMCE-based cassette-exchange systems in animal cell biotechnology," *Cytotechnology*, 2006.
- [226] W. J. Huh, I. U. Mysorekar, and J. C. Mills, "Inducible activation of Cre recombinase in adult mice causes gastric epithelial atrophy, metaplasia, and regenerative changes in the absence of 'floxed' alleles.," *Am. J. Physiol. Gastrointest. Liver Physiol.*, vol. 299, no. 2, pp. G368–80, Aug. 2010.
- [227] E. E. Schmidt, D. S. Taylor, J. R. Prigge, S. Barnett, and M. R. Capecchi, "Illegitimate Cre-dependent chromosome rearrangements in transgenic mouse spermatids.," *Proc. Natl. Acad. Sci. U.S.A.*, vol. 97, no. 25, pp. 13702–13707, Dec. 2000.
- [228] H. J. Gross, B. Verwer, D. Houck, and D. Recktenwald, "Detection of rare cells at a frequency of one per million by flow cytometry," *Cytometry*, 1993.
- [229] A. Rodolosse, A. Barbat, and I. Chantret, "Selecting agent hygromycin B alters expression of glucose-regulated genes in transfected Caco-2 cells," *American Journal of ...*, 1998.
- [230] J. Qiao, A. Oumard, W. Wegloehner, and J. Bode, "Novel tag-and-exchange (RMCE) strategies generate master cell clones with predictable and stable transgene expression properties," *J. Mol. Biol.*, 2009.
- [231] W. L. Kaufman, I. Kocman, V. Agrawal, H.-P. Rahn, D. Besser, and M. Gossen, "Homogeneity and persistence of transgene expression by omitting antibiotic selection in cell line isolation.," *Nucleic Acids Res.*, vol. 36, no. 17, p. e111, Oct. 2008.

- [232] C. Marie, G. Vandermeulen, M. Quiviger, M. Richard, V. Pr  at, and D. Scherman, "pFARs, plasmids free of antibiotic resistance markers, display high-level transgene expression in muscle, skin and tumour cells.," *J Gene Med*, vol. 12, no. 4, pp. 323–332, Apr. 2010.
- [233] R. W. Siegel, R. Jain, and A. Bradbury, "Using an in vivo phagemid system to identify non-compatible loxP sequences," *FEBS Lett.*, 2001.
- [234] J. Bode, T. Schlake, M. Iber, and D. Sch  beler, "The Transgeneticists Toolbox: Novel Methods for the Targeted Modification of Eukaryotic Genomes," *Biological ...*, 2005.
- [235] W. M. Saltzman and D. Luo, "Synthetic DNA delivery systems : Abstract : Nature Biotechnology," *Nat. Biotechnol.*, vol. 18, no. 1, pp. 33–37, Jan. 2000.
- [236] A. Coonrod, F. Q. Li, and M. Horwitz, "On the mechanism of DNA transfection: efficient gene transfer without viruses.," *Gene Ther.*, 1997.
- [237] W. S. Meaking, J. Edgerton, C. W. Wharton, and R. A. Meldrum, "Electroporation-induced damage in mammalian cell DNA.," *Biochim. Biophys. Acta*, vol. 1264, no. 3, pp. 357–362, Dec. 1995.
- [238] J. A. Nickoloff, L. N. Spirio, and R. J. Reynolds, "A comparison of calcium phosphate coprecipitation and electroporation. Implications for studies on the genetic effects of DNA damage.," *Mol. Biotechnol.*, vol. 10, no. 2, pp. 93–101, Oct. 1998.
- [239] M. El‐‐sabhy, A. Nazarali, and M. Foldvari, "Non-viral nucleic acid delivery: key challenges and future directions.," *Curr Drug Deliv.*, vol. 8, no. 3, pp. 235–244, May 2011.
- [240] D. A. Sorrell, C. J. Robinson, J.-A. Smith, and A. F. Kolb, "Recombinase mediated cassette exchange into genomic targets using an adenovirus vector.," *Nucleic Acids Res.*, vol. 38, no. 11, p. e123, Jun. 2010.
- [241] S. R. Braam, C. Denning, S. van den Brink, P. Kats, R. Hochstenbach, R. Passier, and C. L. Mummery, "Improved genetic manipulation of human embryonic stem cells.," *Nat. Methods*, vol. 5, no. 5, pp. 389–392, May 2008.
- [242] D. P. Silver and D. M. Livingston, "Self-excising retroviral vectors encoding the Cre recombinase overcome Cre-mediated cellular toxicity.," *Mol. Cell*, vol. 8, no. 1, pp. 233–243, Jul. 2001.
- [243] A. Loonstra, M. Vooijs, H. B. Beverloo, B. A. Allak, E. van Drunen, R. Kanaar, A. Berns, and J. Jonkers, "Growth inhibition and DNA damage induced by Cre recombinase in mammalian cells.," *Proc. Natl. Acad. Sci. U.S.A.*, vol. 98, no. 16, pp. 9209–9214, Jul. 2001.
- [244] L. Nolden, F. Edenhofer, S. Haupt, P. Koch, F. T. Wunderlich, H. Siemen, and O. Br  stle, "Site-specific recombination in human embryonic stem cells induced by cell-permeant Cre recombinase.," *Nat. Methods*, vol. 3, no. 6, pp. 461–467, Jun. 2006.
- [245] S. Goetze, A. Baer, S. Winkelman, K. Nehlsen, J. Seibler, K. Maass, and J. Bode, "Performance of genomic bordering elements at predefined genomic loci.," *Mol. Cell. Biol.*, vol. 25, no. 6, pp. 2260–2272, Mar. 2005.
- [246] A. Baer, D. Sch  beler, and J. Bode, "Transcriptional properties of genomic transgene integration sites marked by electroporation or retroviral infection.," *Biochemistry*, vol. 39, no. 24, pp. 7041–7049, Jun. 2000.
- [247] M. Vooijs, J. Jonkers, and A. Berns, "A highly efficient ligand-regulated Cre recombinase mouse line shows that LoxP recombination is position dependent.," *EMBO Rep.*, vol. 2, no. 4, pp. 292–297, Apr. 2001.
- [248] D. Hockemeyer and R. Jaenisch, "Gene targeting in human pluripotent cells.," *Cold Spring Harb. Symp. Quant. Biol.*, vol. 75, pp. 201–209, 2010.
- [249] L. Pelkmans, "Cell Biology. Using cell-to-cell variability--a new era in molecular biology.," *Science*, vol. 336, no. 6080, pp. 425–426, Apr. 2012.
- [250] R. M. Kumar and J. J. Collins, "Making a Noisy Gene: HDACs Turn Up the Static.," *Mol. Cell*, vol. 47, no. 2, pp. 151–153, Jul. 2012.
- [251] Y. Kameyama, Y. Kawabe, A. Ito, and M. Kamihira, "An accumulative site-specific gene integration system using Cre recombinase-mediated cassette exchange.," *Biotechnol. Bioeng.*, vol. 105, no. 6, pp. 1106–1114, Apr. 2010.
- [252] E. P. Papapetrou, G. Lee, N. Malani, M. Setty, I. Riviere, L. M. S. Tirunagari, K. Kadota, S. L. Roth, P. Giardina, A. Viale, C. Leslie, F. D. Bushman, L. Studer, and M. Sadelain, "Genomic safe harbors permit high  $\beta$ -globin transgene expression in thalassemia induced pluripotent stem cells.," *Nat. Biotechnol.*, vol. 29, no. 1, pp. 73–78, Jan. 2011.

- [253] R. Lu, N. F. Neff, S. R. Quake, and I. L. Weissman, "Tracking single hematopoietic stem cells in vivo using high-throughput sequencing in conjunction with viral genetic barcoding," *Nat. Biotechnol.*, vol. 29, no. 10, pp. 928–933, Oct. 2011.
- [254] H. Zhou, S. Wu, J. Y. Joo, S. Zhu, D. W. Han, T. Lin, S. Trauger, G. Bien, S. Yao, Y. Zhu, G. Siuzdak, H. R. Schöler, L. Duan, and S. Ding, "Generation of induced pluripotent stem cells using recombinant proteins," *Cell Stem Cell*, vol. 4, no. 5, pp. 381–384, May 2009.
- [255] K. Kaji, K. Norrby, A. Paca, M. Mileikovsky, P. Mohseni, and K. Woltjen, "Virus-free induction of pluripotency and subsequent excision of reprogramming factors," *Nature*, vol. 458, no. 7239, pp. 771–775, Apr. 2009.
- [256] C. M. Swanson and M. H. Malim, "SnapShot: HIV-1 proteins," *Cell*, vol. 133, no. 4, pp. 742–742.e1, May 2008.
- [257] A. C. Logan, S. J. Nightingale, D. L. Haas, G. J. Cho, K. A. Pepper, and D. B. Kohn, "Factors influencing the titer and infectivity of lentiviral vectors," *Hum. Gene Ther.*, vol. 15, no. 10, pp. 976–988, Oct. 2004.
- [258] C. Frecha, C. Costa, D. Nègre, E. Gauthier, S. J. Russell, F.-L. Cosset, and E. Verhoeven, "Stable transduction of quiescent T cells without induction of cycle progression by a novel lentiviral vector pseudotyped with measles virus glycoproteins," *Blood*, vol. 112, no. 13, pp. 4843–4852, Dec. 2008.
- [259] P. de Felipe, "Skipping the co-expression problem: the new 2A 'CHYSEL' technology," *Genet Vaccines Ther.*, vol. 2, no. 1, p. 13, Sep. 2004.
- [260] R. B. Batchu and P. L. Hermonat, "The trans-inhibitory Rep78 protein of adeno-associated virus binds to TAR region DNA of the human immunodeficiency virus type 1 long terminal repeat," *FEBS Lett.*, vol. 367, no. 3, pp. 267–271, Jul. 1995.
- [261] B. A. Antoni, A. B. Rabson, I. L. Miller, J. P. Trempe, N. Chejanovsky, and B. J. Carter, "Adeno-associated virus Rep protein inhibits human immunodeficiency virus type 1 production in human cells," *J. Virol.*, vol. 65, no. 1, pp. 396–404, Jan. 1991.
- [262] N. A. Kokorina, A. D. Santin, C. Li, and P. L. Hermonat, "Involvement of protein-DNA interaction in adeno-associated virus Rep78-mediated inhibition of HIV-1," *J. Hum. Virol.*, vol. 1, no. 7, pp. 441–450, Nov. 1998.
- [263] S. E. Howden, L. Voullaire, and J. Vadolas, "The transient expression of mRNA coding for Rep protein from AAV facilitates targeted plasmid integration," *J Gene Med*, vol. 10, no. 1, pp. 42–50, Jan. 2008.
- [264] Y.-B. Yue, Y.-Y. Xue, L. Tian, J.-L. Xue, J.-Z. Chen, and W. Jia, "Functional differentiation between Rep-mediated site-specific integration and transcriptional repression of the adeno-associated viral p5 promoter," *Hum. Gene Ther.*, vol. 21, no. 6, pp. 728–738, Jun. 2010.
- [265] G. P. van Nierop, A. A. F. de Vries, M. Holkers, K. R. Vrijssen, and M. A. F. V. Gonçalves, "Stimulation of homology-directed gene targeting at an endogenous human locus by a nicking endonuclease," *Nucleic Acids Res.*, vol. 37, no. 17, pp. 5725–5736, Sep. 2009.
- [266] M. A. F. V. Gonçalves, M. Holkers, G. P. van Nierop, R. Wieringa, M. G. Pau, and A. A. F. de Vries, "Targeted chromosomal insertion of large DNA into the human genome by a fiber-modified high-capacity adenovirus-based vector system," *PLoS ONE*, vol. 3, no. 8, p. e3084, 2008.
- [267] M. A. F. V. Gonçalves, M. Holkers, C. Cudré-Mauroux, G. P. van Nierop, S. Knaän-Shanzer, I. van der Velde, D. Valerio, and A. A. F. de Vries, "Transduction of myogenic cells by retargeted dual high-capacity hybrid viral vectors: robust dystrophin synthesis in duchenne muscular dystrophy muscle cells," *Mol. Ther.*, vol. 13, no. 5, pp. 976–986, May 2006.
- [268] M. A. F. V. Gonçalves, G. P. van Nierop, M. R. Tijssen, P. Lefesvre, S. Knaän-Shanzer, I. van der Velde, D. W. van Bekkum, D. Valerio, and A. A. F. de Vries, "Transfer of the full-length dystrophin-coding sequence into muscle cells by a dual high-capacity hybrid viral vector with site-specific integration ability," *J. Virol.*, vol. 79, no. 5, pp. 3146–3162, Mar. 2005.
- [269] A. Giudice and A. Trounson, "Genetic modification of human embryonic stem cells for derivation of target cells," *Cell Stem Cell*, vol. 2, no. 5, pp. 422–433, May 2008.
- [270] Y.-H. Loh, J. C. Yang, A. De Los Angeles, C. Guo, A. Cherry, D. J. Rossi, I.-H. Park, and G. Q. Daley, "Excision of a viral reprogramming cassette by delivery of synthetic cre mRNA," *Curr Protoc Stem Cell Biol*, vol. 4, p. Unit4A.5, 2012.
- [271] M. Conese, C. Auriche, and F. Ascenzioni, "Gene therapy progress and prospects: episomally maintained self-replicating systems," *Gene Ther.*, vol. 11, no. 24, pp. 1735–1741, Dec. 2004.

- [272] U. Moens, M. Van Ghelue, and M. Johannessen, "Oncogenic potentials of the human polyomavirus regulatory proteins," *Cell. Mol. Life Sci.*, vol. 64, no. 13, pp. 1656–1678, Jul. 2007.
- [273] S. Humme, G. Reisbach, R. Feederle, H.-J. Delecluse, K. Bousset, W. Hammerschmidt, and A. Schepers, "The EBV nuclear antigen 1 (EBNA1) enhances B cell immortalization several thousandfold," *Proc. Natl. Acad. Sci. U.S.A.*, vol. 100, no. 19, pp. 10989–10994, Sep. 2003.
- [274] M. M. P. Lufino, R. Manservigi, and R. Wade-Martins, "An S/MAR-based infectious episomal genomic DNA expression vector provides long-term regulated functional complementation of LDLR deficiency," *Nucleic Acids Res.*, vol. 35, no. 15, p. e98, 2007.
- [275] D. M. Gilbert and S. N. Cohen, "Bovine papilloma virus plasmids replicate randomly in mouse fibroblasts throughout S phase of the cell cycle," *Cell*, vol. 50, no. 1, pp. 59–68, Jul. 1987.
- [276] R. E. White, R. Wade-Martins, S. L. Hart, J. Frampton, B. Huey, A. Desai-Mehta, K. M. Cerosaletti, P. Concannon, and M. R. James, "Functional delivery of large genomic DNA to human cells with a peptide-lipid vector," *J Gene Med*, vol. 5, no. 10, pp. 883–892, Oct. 2003.
- [277] M. Kakeda, M. Hiratsuka, K. Nagata, Y. Kuroiwa, M. Kakitani, M. Katoh, M. Oshimura, and K. Tomizuka, "Human artificial chromosome (HAC) vector provides long-term therapeutic transgene expression in normal human primary fibroblasts," *Gene Ther*, vol. 12, no. 10, pp. 852–856, May 2005.
- [278] M. Katoh, F. Ayabe, S. Norikane, T. Okada, H. Masumoto, S.-I. Horike, Y. Shirayoshi, and M. Oshimura, "Construction of a novel human artificial chromosome vector for gene delivery," *Biochem. Biophys. Res. Commun.*, vol. 321, no. 2, pp. 280–290, Aug. 2004.
- [279] C. Magin-Lachmann, G. Kotzamanis, L. D'Aiuto, H. Cooke, C. Huxley, and E. Wagner, "In vitro and in vivo delivery of intact BAC DNA -- comparison of different methods," *J Gene Med*, vol. 6, no. 2, pp. 195–209, Feb. 2004.
- [280] J. Gómez-Rodríguez, V. Washington, J. Cheng, A. Dutra, E. Pak, P. Liu, D. W. McVicar, and P. L. Schwartzberg, "Advantages of q-PCR as a method of screening for gene targeting in mammalian cells using conventional and whole BAC-based constructs," *Nucleic Acids Res.*, vol. 36, no. 18, p. e117, Oct. 2008.
- [281] P. de Felipe, G. A. Luke, L. E. Hughes, D. Gani, C. Halpin, and M. D. Ryan, "E unum pluribus: multiple proteins from a self-processing polyprotein," *Trends in Biotechnology*, vol. 24, no. 2, pp. 68–75, Feb. 2006.
- [282] J. Sambrook and D. W. Russell, "Calcium-phosphate-mediated Transfection of Eukaryotic Cells with Plasmid DNAs," *CSH Protoc*, vol. 2006, no. 1, 2006.
- [283] M. Glogauer and C. A. McCulloch, "Introduction of large molecules into viable fibroblasts by electroporation: optimization of loading and identification of labeled cellular compartments," *Exp. Cell Res.*, vol. 200, no. 2, pp. 227–234, Jun. 1992.
- [284] K. A. Page, N. R. Landau, and D. R. Littman, "Construction and use of a human immunodeficiency virus vector for analysis of virus infectivity," *J. Virol.*, 1990.
- [285] G. Tiscornia, O. Singer, and I. M. Verma, "Production and purification of lentiviral vectors," *Nature protocols*, 2006.
- [286] M. Geraerts, S. Willems, V. Baekelandt, Z. Debyser, and R. Gijssbers, "Comparison of lentiviral vector titration methods," *BMC Biotechnol.*, vol. 6, p. 34, 2006.
- [287] D. Trono, "Partial reverse transcripts in virions from human immunodeficiency and murine leukemia viruses," *J. Virol.*, vol. 66, no. 8, pp. 4893–4900, Aug. 1992.
- [288] C. Delenda and C. Gaillard, "Real-time quantitative PCR for the design of lentiviral vector analytical assays," *Gene Ther*, vol. 12, pp. S36–50, Oct. 2005.
- [289] P. R. Langer, A. A. Waldrop, and D. C. Ward, "Enzymatic synthesis of biotin-labeled polynucleotides: novel nucleic acid affinity probes," *Proc. Natl. Acad. Sci. U.S.A.*, vol. 78, no. 11, pp. 6633–6637, Nov. 1981.